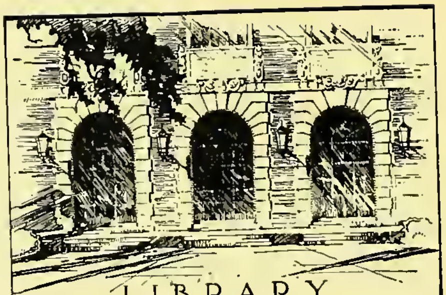


TRAINING RESEARCH LABORATORY

BUREAU OF EDUCATIONAL RESEARCH
DEPARTMENT OF PSYCHOLOGY
University of Illinois
Urbana, Illinois

LEARNING HOW TO LEARN UNDER SEVERAL CUE CONDITIONS

DALE E. MATTSON

TECHNICAL REPORT NO. 1
AUGUST, 1963

PSYCHOLOGICAL AND EDUCATIONAL FACTORS IN
TRANSFER OF TRAINING
Phase I

U. S. OFFICE OF EDUCATION
Contract 2-20-003

Lawrence M. Stolurow
Principal Investigator

TRAINING RESEARCH LABORATORY
University of Illinois
Urbana, Illinois

Learning How to Learn Under Several Cue Conditions[1]

Dale E. Mattson

Technical Report No. 1

August 1963

Psychological and Educational Factors in Transfer of Training
Phase I

Principal Investigator:

Lawrence M. Stolurow

Project Sponsor:

Educational Media Branch
U.S. Office of Education
Title VII
Contract #2-20-003

---

[1]Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Education in the Graduate College of the University of Illinois, 1963.

LEARNING HOW TO LEARN UNDER SEVERAL CUE CONDITIONS

Dale Edward Mattson, Ph.D.

Department of Education

University of Illinois, 1963

The first objective of this experimental study was to identify and compare three kinds of transfer effects: an effect associated with cue repetition, a learning-to-learn effect, and a warm-up effect. The second major objective was to evaluate the usefulness of cue-response criterialities in explaining transfer effects.

A factorial design was employed with three degrees of similarity between the relevant cues for the training tasks and those for the criterion task and two degrees of similarity between type of training task and criterion task.

Conditions of Cue Similarity. For one third of the experimental Ss relevant and irrelevant cues remained the same for all tasks; for another third relevant and irrelevant cues were reversed on the criterion task; and for the remaining third completely new cues were introduced during the criterion task.

Conditions of Task Similarity. For half of the experimental Ss training and criterion tasks were of the same type. For the other half training tasks and criterion tasks were quite different.

In addition to the six experimental groups necessary for the experimental design an additional group of Ss was a control group who performed only the criterion task.

The entire experiment was carried out twice--once using large group testing procedures and once testing groups of either 7 or 14 at a time.

The Ss for this experiment were undergraduate college students. For the first experiment in which large group testing procedures were used the Ss participated in the experiment as part of a course requirement either in introductory psychology or in educational psychology. For the second experiment all Ss volunteered.

The results of the study may be summarized as follows:

1. Ss of the three cue conditions did not differ significantly on the number of errors made during the completion of the criterion task.

2. A learning-to-learn effect was identified. Ss who received training on a series of training tasks similar to the criterion task completed the criterion task with fewer errors than Ss for whom training tasks were unlike the criterion task.

3. A warm-up effect was identified. Subjects who performed a series of four tasks quite different from the criterion task, using cues unlike those used on the criterion task, completed the criterion task with fewer errors than Ss in the control group.

4. Using the same two cues in the solution of a number of training tasks increased the use of these cues on the first trial of the criterion task. The criterialities (correlations between cues and responses) were higher on the first trial of the criterion task for cues that had previously been relevant than for cues that had been irrelevant.

ACKNOLWEDGMENT

The author wishes to express his special thanks to committee chairman, Dr. L. M. Stolurow, who has directed all of the author's graduate work at the University of Illinois, and to Dr. L. J. Cronbach who as acting chairman has supervised the final preparation of this dissertation. Helpful suggestions given by other members of the committee, Dr. J. A. Adams, Dr. R. S. Jones, and Dr. R. E. Pingry are also appreciated.

Special mention should be made of the help given by Robert Yackel of the Statistical Service Unit in processing the IBM cards.

Grateful scknowledgment is made of the financial aid given through a grant from the U. S. Office of Education, Title VII, Project No. 2-20-003, which helped make the study possible.

The author also expresses his sincere gratitude for fellowship funds provided by the U. S. Office of Education, Title 4.

Finally, sincere appreciation is given to Mary Mattson for help and inspiration through all phases of the project.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Transfer effects may be defined as changes in ability to deal with situations not encountered during training (Cronbach, 1963). The range of conditions that lead to transfer may be illustrated by studies of the learning of lists of paired-associates. In an early experiment Bruce (1933) studied the transfer resulting under several conditions of similarity between training lists and the criterion list. The greatest positive transfer occurred when the response terms were identical for training lists and the criterion list. However, positive transfer also occurred when the lists used for training were completely different from the criterion list. In a more recent study Thune (1950) demonstrated that color guessing had a facilitating effect on the subsequent learning of a list of paired-associates.

As we consider these examples in turn, the similarity between training task and the criterion task becomes progressively smaller. For Bruce's first condition the similarity between the training task and the criterion task is evident. For his second condition the similarity is much less obvious; only the requirements of the training tasks and the criterion task are the same. Finally, it is difficult to identify any similarity between Thune's training task (color guessing) and his criterion task (learning a list of paired-associates).

Mandler (1962) has suggested that three kinds of transfer effects can be distinguished, depending on the kind of similarity which exists between training tasks and the transfer task. The first of these is

a transfer effect based on overlearning of training tasks in an A-B, A-C experimental design,--hence, where there is cue repetition. The second is a learning-to-learn effect, and the third is a warm-up effect.

Mandler was primarily interested in the stage of learning of training tasks in the A-B, A-C design. Other variables of interest might be similarity of response, time interval between tasks, etc. In order to include all studies involving the A-B, A-C design, Mandler's first kind of transfer will be referred to in this paper as transfer based on cue repetition.

Tasks used in studying transfer effects based on cue repetition often require Ss to learn which cues in a stimulus situation are relevant. A number of studies demonstrate that when a cue is found to be relevant for one task it is more likely to be regarded as relevant for further tasks (Eckstrand and Wickens, 1954; Lawrence, 1949, 1950; and Stolurow and Solley, 1955). When the old cue is indeed relevant for the new task positive transfer is observed. If, however, the old cue is no longer relevant (a nonreversal shift) negative transfer is observed (Harrow and Friedman, 1958; and Kendler and D'Amato, 1955). A summary of effects of nonreversal shifts are summarized in Concept Learning by Hunt (1963, Pp74-78).

The second kind of transfer effect suggested by Mandler is a learning-to-learn effect. The phrase "learning-to-learn" has a number of possible meanings. For example, in popular jargon it might be said that a person is learning to learn when he takes a course entitled "How to study effectively." Here a person might be expected to learn some study techniques which would be useful in a wide variety of tasks.

At the other extreme, Ss in a psychology experiment may learn a very specific solution-rule which is only applicable to a certain type of problem. Harlow has referred to this type of learning to learn as the formation of learning sets (1949, 1960).

In the present paper Mandler's definition of learning to learn will be used. He speaks of learning to learn as the facilitation which occurs when Ss are given a series of repeated and related tasks. Therefore, learning to learn would include both facilitation due to the learning of a solution rule, and facilitation due to the learning of more general techniques for improving performance. Facilitation which occurred when Harlow's (1944, 1960) chimps learned an "oddity principle" is an example of learning to learn which involves the discovery of a solution-rule. An example of learning to learn in which Ss learn more general techniques for improving performance may be found in Thune's (1951) study in which Ss learned a series of lists of paired-associates. In addition to learning individual S-R associations, Ss evidently learned some general techniques which helped them in the learning of further lists.

Although Harlow's studies are perhaps the best known studies of learning to learn, a number of other studies have shown similar results. The facilitation resulting from the learning of completely separate lists of paired-associates in the study by Bruce (1933) and transfer from one psychomotor task to another in a study by Cox (1933) are early examples of learning to learn. More recent studies by Adams (1954), Duncan (1958, 1960), Shepard (1957), and Thune and Eriksen (1960) have also demonstrated positive transfer resulting from practice on tasks related to, but unlike, the criterion task.

The third kind of transfer effect suggested by Mandler is a warm-up effect. Mandler credits Irion (1948) with first identifying this type of transfer. This is a short-term transfer effect which occurs within a particular experimental period. It may be due to attention habits, reduced tension, etc. The facilitation due to color guessing in Thune's experiment (1950) would come under this classification, as would transfer effects obtained in studies by Hamilton (1950) and Mandler (1956). A complete review of the warm-up literature may be found in a recent article by Adams (1961).

Mandler points out the necessity of using control groups in distinguishing between the various kinds of transfer effects. He claims that this has not been done in most studies so that "warm-up and learning set effects...are usually confounded. Thus it is often not determinable to what extent an animal's prior experience in a maze produces varying degrees of specific postural and attentive habits (warm-up) as against non-specific structural effects (learning set)" (1962, Pp. 421). When Thune (1950) attempted to distinguish between warm-up effects and learning-to-learn effects he failed to find any significant differences in favor of his learning-to-learn group.

In most studies of transfer effects based on cue repetition, these effects are confounded both with learning to learn effects and with warm-up effects (D'Amato and Jagoda, 1960; Harrow and Friedman, 1958; Kendler and D'Amato, 1955; and Kendler and Kendler, 1958). In a study by Kelleher (1956) a control group was used in order to distinguish between transfer due to cue repetition and transfer due to learning to learn. All Ss including Control Ss received training on a discrimination task similar

to the test task. Group differences during the performance of the test task were therefore attributable solely to conditions of cue similarity between training and test tasks. It was not possible to compare learning-to-learn effects with cue effects because no group performed only the test task.

Thus although the three kinds of transfer suggested by Mandler have been previously demonstrated, there is very little information available as to the relative size of each. One of the major objectives of the present study was, therefore, to compare these three kinds of transfer in one experimental setting.

The second major purpose of this study was to evaluate the usefulness of a novel method of analyzing data. In most transfer studies, transfer scores are used to make inferences about the mediational processes of Ss. For example in the study by Eckstrand and Wickens (1954) the following conclusion was reached. "It may be inferred from the performance on the test tasks (number of trials to criterion) that the prior experience with the relevance and irrelevance of certain dimensions on the first two tasks developed biases which influenced the predominating cue on the third task." It would be more satisfactory if mediational processes of Ss could be identified objectively so that these mediational processes could be used in interpreting transfer effects.

Brunswik (1956) suggested that the learning of mediational processes could be observed in the emergence of correlations between cues and the responses of the Ss. Bruner, Goodnow, and Austin (1956) termed this cue-response correlation "degree of criteriality." Bruner *et al.* describe cue criteriality in the following manner. "Take the category

of things called 'apples' by some particular person. We are interested in those attributes that affect the probability of our person calling an object an apple. Insofar as changes in the values of any particular attribute do not produce changes in the probability of the object being called an apple, we call that attribute noncriterial. Any attribute which when changed in value alters the likelihood of an object being categorized in a certain way is, therefore, a criterial attribute for the person doing the categorizing." (Bruner *et al.*, 1956).

In a number of studies an attempt has been made to identify mediational processes by means of cue criterialities (Azuma, 1960; Cronbach and Azuma, 1961b; McHale and Stolurow, 1962; and Smedslund, 1955). In all of these studies *Ss* were expected to learn to make scaled responses to displays containing scaled cues. For each *S* product-moment correlations were computed between the values of cues and the responses made. These correlations (criterialities) were computed over blocks of overlapping trials for each person separately. As predicted, average criterialities for relevant cues tended to rise to the values which would indicate ideal weighting and average criterialities for nonrelevant cues approached zero.

Criterialities computed over blocks of trials are not altogether satisfactory as a means of learning about the mediational processes of *Ss*. If *S* develops a classification system for stimulus displays and then responds differently to different classes of displays, this information is lost in criterialities computed over blocks of many trials. According to Cronbach and Azuma (1961a) this seems to be the way *Ss* viewed the Azuma problem. They concluded that "computing criterialities

was a fine way to analyze data under the Brunswikian hypothesis that S responds to an aggregate of stimuli--but is there any objective procedure for inferring how an S forms categories and hypotheses applied with categories?"

There are at least two other situations in which criterialities computed over blocks of trials would not be very useful. Suppose that one wished to study the learning which takes place in a task requiring only five or six trials to a solution. Criterialities computed over a small block of trials would be too unreliable to be very useful. Another situation in which block criterialities are not very useful may be found in studies of transfer effects. For example, suppose that a study were designed to identify transfer effects based on the repetition of relevant cues in two consecutive tasks. In this case one would be interested in demonstrating an increased criteriality for these previously relevant cues on the initial trial of the transfer task, not over a block of trials.

What is needed is a method of computing the criteriality of a cue for a single trial. Although this cannot be done for one individual, it should be possible to compute the criteriality of a cue for a group of individuals on a single trial by presenting each individual with a different display. By this procedure one could compute the single-trial criteriality for each cue for the first trial of a transfer task. This procedure will be followed in this study in order to identify the effects of performing a series of tasks in which the same cues are present.

The two objectives of this study may be translated into the following hypotheses:

1. Three kinds of transfer effects can be identified and compared: an effect associated with cue repetition, a learning-to-learn effect, and a warm-up effect.
2. Cue repetition is expected to result in a negative effect under a condition similar to a nonreversal shift (relevant cues during training become irrelevant during the criterion task) and a positive effect under a condition in which the same cues are relevant for training task and criterion task.
3. On the first trial of the transfer task, single-trial criterialities will be higher for cues previously relevant than for cues previously irrelevant.

CHAPTER II

METHOD

Subjects and Testing Schedule

Altogether, 240 undergraduate college students were used in this study. Group testing procedures were used. The experiment was carried out twice; once using large groups and once using groups of either 7 or 14. The experimental design required the use of 112 Ss.

In the first experiment it was planned to test all Ss at the same time. Due to an error in scheduling, a room large enough to seat only 140 people was assigned for the experiment. When 101 Ss appeared it was necessary for Ss to sit next to each other. Fourteen of these Ss performed in such a way that their data were not usable. Either they failed to follow directions or they failed to complete the training tasks satisfactorily. In order to complete the data for this first experiment, 25 further Ss were tested.

Because of the necessity of replacing so many Ss and because of the adverse testing conditions during the first experiment, the entire experiment was repeated using smaller groups. Within each small group an equal number of Ss was assigned to each experimental condition and to the control group. Two Ss were dropped from the second experiment; one for failing to follow directions and the other for failing to complete the training tasks satisfactorily.

All of the Ss for the first experiment were required to take part in the experiment as part of a course requirement either for introductory psychology or for educational psychology. All of the Ss for the second experiment were volunteers.

Description of Tasks

Two types of problems or tasks were used in this study. The first of these was patterned after that of Azuma (1960). For this task, S had to learn to make scaled responses by weighting cues; the first type of task will therefore be referred to as a "W" (weighting) task. The second type of task was a conjunctive concept formation task (Bruner et al., 1956) in which a concept was defined by the presence or absence of X's inside two closed figures. This type of task will be referred to as an "X" task.

Type W Tasks. Figure 1 contains three typical stimulus displays used for type W tasks. In order to complete each task S had to learn to make a correct numerical response to each of a series of such displays. After S responded to each display, feedback was given by allowing him to see the correct answer. Correct answers for all displays could be determined by using a formula; the formula changed from task to task. The correct formula for one W task, for example, was as follows: multiply the numbers in the square and the circle by two and one respectively and then add. Using this formula, the correct answers to the three stimulus displays in Figure 1 would be 4, 5, and 6.

The numbers inside only two of the four figures in each display were relevant and these numbers were weighted by sets of constants. The sets of constants for the five W tasks were as follows: 1,1; 2,2; 2,1; 1,3; and 1,$\frac{1}{2}$. Information as to which figures contained relevant numbers for each S and for each task will be given in a later section.

Type X Tasks. Figure 2 contains three typical stimulus displays used for type X tasks. S had to learn to label displays as "K" displays

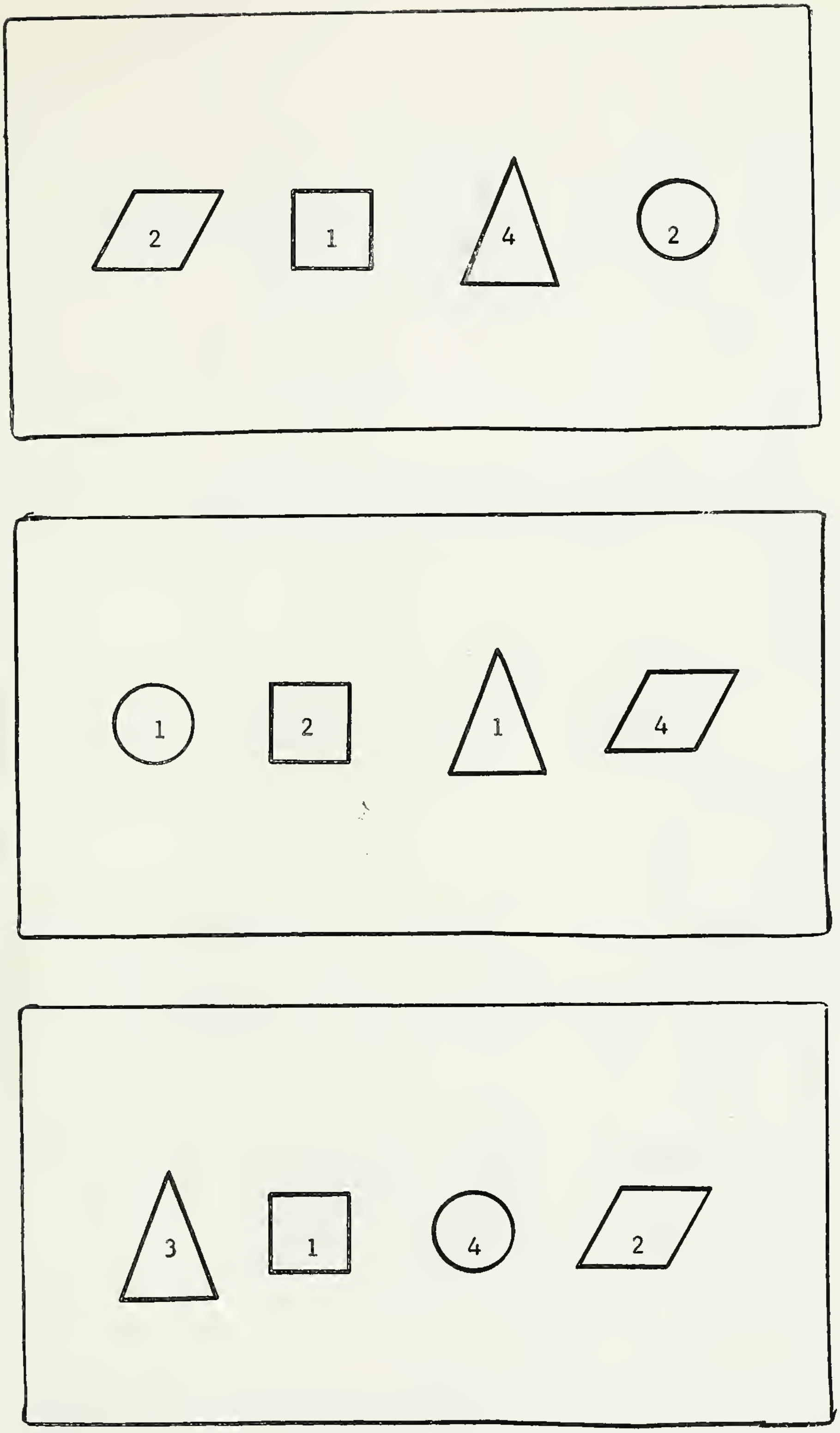


FIGURE 1

TYPICAL STIMULUS DISPLAYS FOR TYPE W TASKS

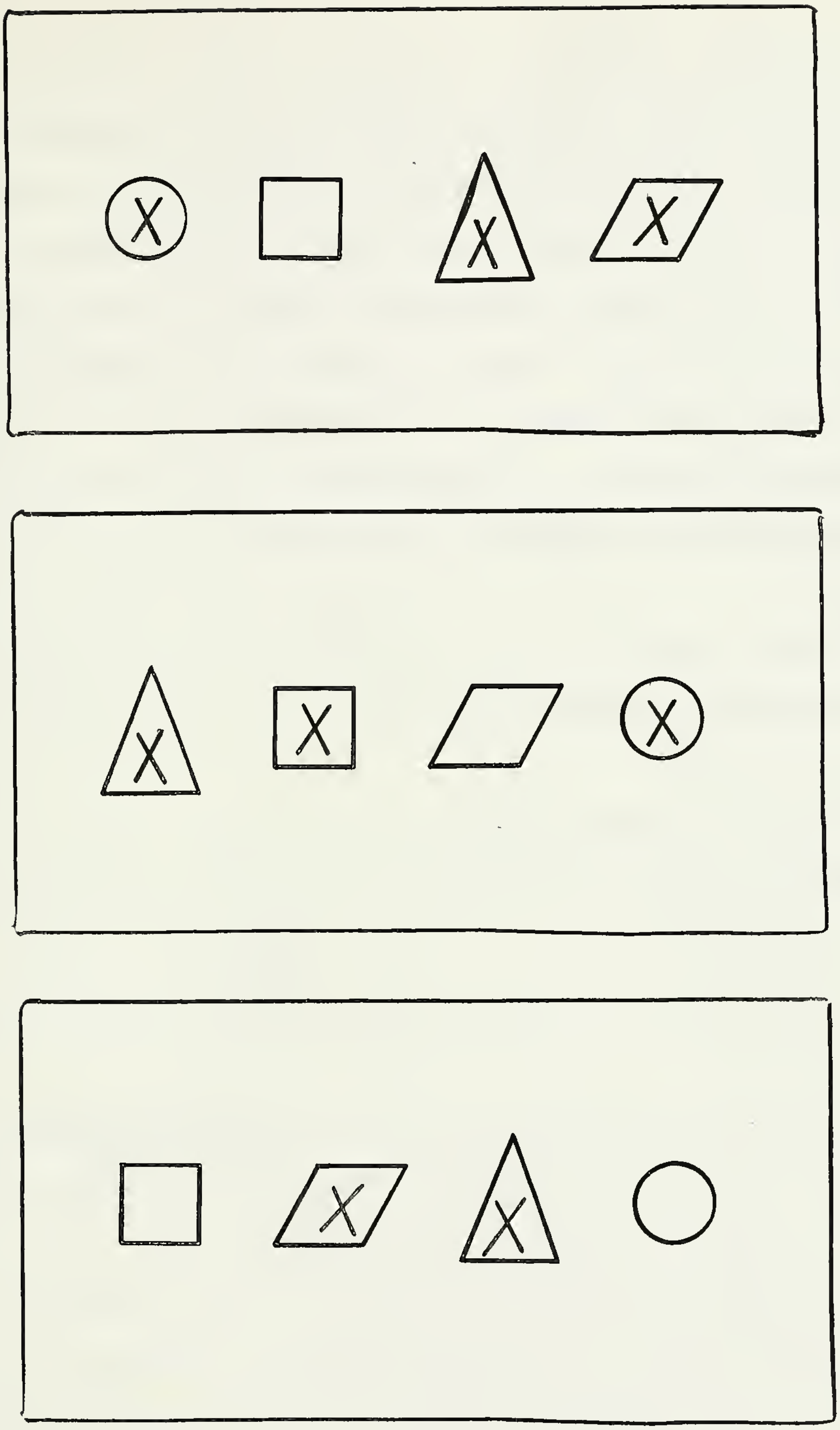


FIGURE 2

TYPICAL STIMULUS DISPLAYS FOR TYPE X TASKS

or as "O" displays. After S responded to each display, feedback was given by allowing him to see the correct answer. For each task correct answers for all displays could be determined by using a single rule. The correct rule for the first type X task, for example was: any display in which the circle and the square contain X's is a "K" display; all others are "O" displays. (Actually relevant figures were counterbalanced as will be explained in a later section. Until then descriptions will be given as though the circle and square were always relevant.) Using this rule the correct answers for the three displays in Figure 2 would be O, K, and O.

For each of the three remaining type X tasks S was required to discover a similar rule. These rules for identifying K displays were as follows:

2nd X task - The circle and the square both had to be empty.

3rd X task - The circle had to contain an X; the square had to be empty.

4th X task - The circle had to be empty; the square had to contain an X.

Termination of Tasks. Studies have shown that Ss generally can distinguish examples of a concept before they can verbalize a correct definition of it (Hull, 1920; Smoke, 1932; Walk, 1952; Adams, 1957; and Davis and Hess, 1962). In order for the required mediational processes to be learned in the present study it was necessary that S become aware which cues were relevant for each of the training tasks. The criterion for the completion of each task was 16 consecutive correct responses accompanied by a correct verbalization of the solution.

Experimental Design

The basic design of this experiment is a factorial design involving two degrees of similarity between training tasks and the criterion task and three degrees of similarity between the cues used for the training tasks and those used for the criterion task (throughout this paper the word "cue" will refer to one of the four closed figures in a stimulus display.) In addition to the six groups (16 Ss per group) necessary for this design, an additional group of Ss was used as a control group. These control Ss performed only the criterion task. Table 1 contains a schematic description of the training and transfer conditions for the seven groups in both experiments.

TABLE 1

SCHEMATIC DESCRIPTION OF TRAINING AND TRANSFER CONDITIONS FOR EACH GROUP

Group	Training Tasks: Relevant Cues	Training Tasks: Irrelevant Cues	Training Tasks: Type of Task	Criterion Task: Relevant Cues	Criterion Task: Irrelevant Cues	Criterion Task: Type of Task
WS	ab	cd	W	ab*	cd*	W*
WO	cd	ab	W	ab**	cd	W*
WN	ef	gh	W	ab	cd	W*
XS	ab	cd	X	ab*	cd*	W
XO	cd	ab	X	ab**	cd	W
XN	ef	gh	X	ab	cd	W
Control	No Training Tasks			ab	cd	W

*Same as in training
**Formerly irrelevant

Conditions of Task Similarity. Two conditions of similarity between training tasks and the criterion task were used in this study. For all Ss the criterion task was a W task with constants 1/2 and 1. Before

performing the criterion task all experimental Ss performed four training tasks. Half of the experimental Ss performed four type W tasks (weights: 1,1; 2,2; 2,1; 1; 3) before performing the criterion task. These Ss will be referred to as W groups. The remaining experimental Ss (X groups) performed four type X tasks before performing the criterion task.

Conditions of Cue Similarity. Three conditions of cue similarity between training and criterion tasks were used. For one third of the experimental Ss the same four cues were present during all tasks and the same two of these four cues were relevant throughout. This is the S (same cue) condition. The S condition is represented in Table 1 by two groups; WS and XS. Ss of the WS group received training on W tasks while Ss of the XS groups received training on X tasks. For example, an XS subject might be trained on tasks where the presence or absence of X's in the circle and square was always significant; in his criterion (W) task the numbers to be weighted appeared in the circle and square.

For another third of the experimental Ss the same four cues were present throughout all tasks but the two cues relevant for training tasks were irrelevant for the criterion task and vice versa (a nonreversal shift). This is the (opposite cue) condition represented in Table 1 by group WO and group XO.

For the remaining experimental Ss four completely new cues, unlike the four present during training task, were introduced for the criterion task. This is the N (new cue) condition. The four cues used during training for the Ss in this third cue condition are shown in Figure 3. For the WN group these figures contained numbers, while for the XN group they were either empty or else they contained X's.

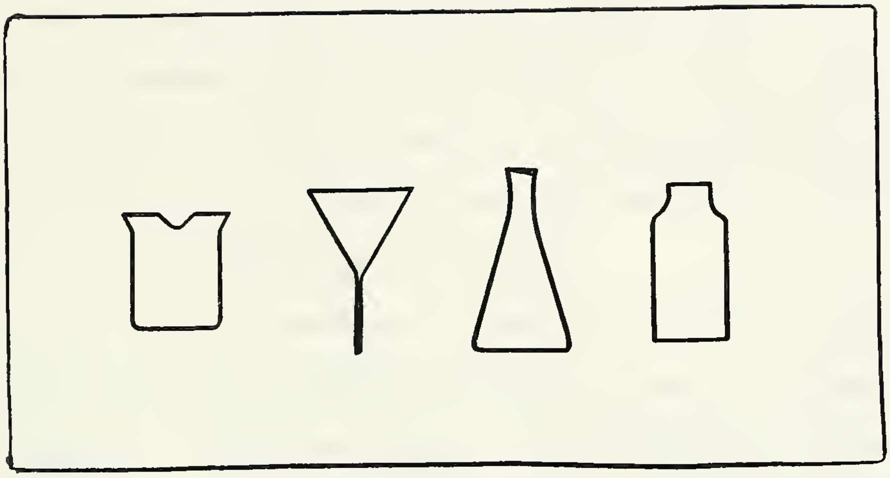


FIGURE 3

CUES USED DURING TRAINING FOR N (NEW CUE) GROUPS

Description of Displays

Order of Cues. The cues did not appear in the same order for all of the 80 displays used for each task. Sixteen of the 24 possible arrangements of the four cues were chosen by E. These orders were then used to make up five blocks of 16 displays each. For a description of these orders see Appendix A.

Symbols within Cues: Type W Tasks. For a study similar to this one McHale and Stolurow (1962) constructed a set of stimulus displays as follows. Within each block of 16 displays all possible combinations of the four values of the two relevant cues occurred once. This made

the correlation between these values zero. The distribution of the values of the two irrelevant cues was rectangular. Correlations between all possible combinations of cues and between values of irrelevant cues and correct answers were all less than .10. Correlations between the correct answers and numbers inside the cues weighted 1/2 and 1 were .45 and .89 respectively. The five blocks of 16 stimulus displays used for type W tasks for this study contained the same numbers as the first five blocks of displays used by McHale and Stolurow.

In order to make it possible to compute single-trial criterialities for any given trial, each S within a group had to see a different stimulus display on that trial. This was accomplished by using the cue values of the first 16 displays employed by McHale and Stolurow so that on the first trial all 16 Ss in a group saw different displays. On the first trial, the first S saw display number one; the second S saw display number two; etc. On the next trial all of the displays shifted one position so that at the end of 16 trials every S had seen the same sequence, but each had begun at a different point. On the 17th trial a new set of 16 displays was introduced.

Symbols within Cues: Type X Tasks. For type X tasks there were 16 possible combinations of filled and empty figures. Four out of these 16 possible stimulus displays are examples of whatever conjunctive concept was to be learned. For each block of 16 displays a new random arrangement was used. See Appendix B for a complete description of the 80 displays used for type X tasks.

Rotation of Cues Over Ss. In order to balance the effects of "Eindringlichkeit" (Brunswik, 1938) or initial cue preference, not all

Ss within any group used the same stimulus figures as relevant cues during the criterion task. The choice of relevant cues from among the four geometric figures (a circle, a square, a triangle, and a rhombus) was balanced within each group. For the first four Ss in each group K = 1/2 (circle) + 1 (square), for the next four Ss K = 1/2 (square) + 1 (triangle), etc. Table 2 contains a complete description of the cues which were relevant and the cues which were irrelevant for each S on the criterion task and on the four training tasks.

Method of Presentation and Instructions to Ss.

Description of Booklets. Booklets with different colored covers were made up for each task. Each booklet contained complete written instructions and sample displays. Answers were recorded on separate answer sheets. Booklets were in the "zebra stripe" form with five displays on each page. S followed the proper sequence of displays by responding to only one display on each page and then turning to the following page for the next display. He would thus go through the entire booklet responding to the first display on each page--the one at the top. Then he would return to the front of the booklet and respond to the second display on each page.

The displays and the correct answers appeared on the same page. The answers were covered by tabs of paper which could be lifted to reveal the correct answer.

Written Instructions to Ss. In general the written instructions (see Appendix C) gave two kinds of information which would be of aid in solving the problems. First, the general rule for the solution to

TABLE 2

ASSIGNMENT OF FIGURES AS RELEVANT AND IRRELEVANT CUES

Groups	Subject Numbers	Training Tasks: Relevant Cues	Training Tasks: Irrelevant Cues	Transfer Tasks: Relevant Cues	Transfer Tasks: Irrelevant Cues
WS or XS	1-4	circle & square	triangle & rhombus	circle & square	triangle & rhombus
	5-8	square & triangle	rhombus & circle	square & triangle	rhombus & circle
	9-12	triangle & rhombus	circle & square	triangle & rhombus	circle & square
	13-16	rhombus & circle	square & triangle	rhombus & circle	square & triangle
WO or XO	1-4	triangle & rhombus	circle & square	circle & square	triangle & rhombus
	5-8	rhombus & circle	square & triangle	square & triangle	rhombus & circle
	9-12	circle & square	triangle & rhombus	triangle & rhombus	circle & square
	13-16	square & triangle	rhombus & circle	rhombus & circle	square & triangle
WN or XN	1-4	beaker & funnel	flask & jar	circle & square	triangle & rhombus
	5-8	beaker & funnel	flask & jar	square & triangle	rhombus & circle
	9-12	beaker & funnel	flask & jar	triangle & rhombus	circle & square
	13-16	beaker & funnel	flask & jar	rhombus & circle	square & triangle
Control	1-4	No Training Tasks		circle & square	triangle & rhombus
	5-8			square & triangle	rhombus & circle
	9-12			triangle & rhombus	circle & square
	13-16			rhombus & circle	square & triangle

the problem was given. Second, sample displays along with correct answers were given for the problem about to be undertaken.

For all type W problems the rule for the solution was explained as follows; "The K-value of a display is affected only by the numbers inside two of the figures and K is obtained by multiplying the numbers inside each of these relevant figures by some constant and then adding." For each problem four sample displays with their correct answers were given. In the first sample display each figure contained a numeral "1", in the second each contained "2", etc. Beneath these sample figures S was told that correct answers to displays for that problem would never be greater or less than certain numbers which were given. This was done in order to restrict the range of responses of Ss.

For type X problems the solution rule was explained as follows. "Only two of the four figures are used in determining correct answers, and it is the presence or absence of X's in these two relevant figures which determines whether a display is a K display or not." Two sample displays and their correct answers were given. In the first sample display all figures contained X's; in the second sample display all figures were empty.

Oral Instructions to Ss. After supplying personal data, Ss of the first experiment were told orally: "You are going to be asked to solve a series of problems. Each booklet contains a single problem. You are to work with the booklets in the following order: white, rust, blue, green, and grey. Some of you have only one booklet. In this case you have only one problem to solve. Now everyone turn to the instructions in the first booklet. You are to study the instructions until I tell

you to stop. You will have seven minutes. Do not go past the page of sample figures."

After seven minutes the following instructions were given: "I want to emphasize six things which were in the directions. Number one, you are to write an answer for every display before looking at the answer even if you are only guessing. Two, you are to answer only one display on each page before proceeding to the next page. Three, only two figures are relevant in any problem. Four, order of figures within a display is completely irrelevant. Five, when you get 16 consecutive correct answers write the rule which you are using at the bottom of the page and then go to the next problem. And finally six, the directions indicated that you would be told when twenty seconds had elapsed so that you could pace yourself. You need not feel forced to keep exact pace with the timer. However you should use the timer as an approximate timing device if you are going to finish these problems in the required time. (For both experiments a tape recorder was used as a timing device. A loop of tape containing the single word "turn" was played through the recorder every 20 seconds.)

"Now, those of you who would like to may spend some more time on the instructions and the rest of you may begin on the first task whenever you are ready."

For the second experiment the only change in the oral instructions was that S was told to have each solution checked before proceeding to the next problem.

In both experiments any S who had failed to solve the first task at the end of 80 trials was told to lift all of the tabs from several pages so that he could examine several displays and answers simultaneously. Under these conditions all Ss solved the first task.

Predicted Results

A number of predictions were made during the planning of this experiment. There were, first, predictions about the relative number of errors each group would make on the criterion task, and second, predictions regarding the criterialities of the several cues during the criterion task.

The following predictions were made about criterion task performance.

1. S (same cue) groups would make fewer errors than N (new cue) groups, while O (opposite cue) groups would make more errors than N groups.
2. W (type W training tasks) groups would make fewer errors than X (type X training tasks) groups.

The first prediction is based on the hypothesis that S groups would show positive transfer due to the repetition of relevant cues and that the O groups would show negative transfer due to the reversal of relevant and irrelevant cues. N groups provided a basis for comparison since neither facilitating nor interfering effects from cue similarity were expected. The second prediction is based on the hypothesis that W groups would exhibit a learning-to-learn effect.

Warm-up effects differ from learning-to-learn effects in two ways: first for warm-up effects there is a lack of similarity between training and criterion tasks, and second, warm-up effects are very sensitive to the time interval between training and criterion tasks (Adams, 1961). In the present experiment the difference between the performances of the XN group and the control group will be referred to as a warm-up effect

on the basis of the lack of similarity between the training tasks and the criterion task for the XN group. The design of the experiment did not include as a variable the time interval between tasks. Such a design would of course be possible. If the effect which has been labeled a warm-up effect is correctly labeled this effect should be sensitive to the length of the interval between training and criterion tasks.

The following predictions were made regarding the single-trial criterialities of the cues on the criterion task:

1. For S and O groups single-trial criterialities for the first trial of the criterion task will be greater for cues relevant on preceding tasks than for cues irrelevant on preceding tasks.
2. For control groups and for N groups there will be no differences larger than chance expectancy among the single-trial criterialities of cues on the first trial of the criterion task.
3. For all groups single-trial criterialities for relevant cues will approach the correct criterialities of .44 and .89 (see section entitled Description of Displays.) For non-relevant cues the single-trial criterialities will approach zero.

RESULTS

Group Differences

Error Scores. The score used as a measure of the rate of learning was the number of errors made by S before reaching the criterion of 16 consecutive correct answers. For each task the minimum error score possible would be zero, for Ss who responded correctly for the first 16 displays; this could occur only with very lucky initial trials. The maximum error score possible would be 80, for Ss who failed to make correct responses for any of the 80 displays. Total error scores were used rather than number of trials to criterion to avoid placing undue emphasis on chance errors in arithmetic. Once S reached criterion, all remaining trials were considered correct. Appendix D contains a complete listing of the number of errors made by each S on each task. Groups means, medians, and standard deviations are also included.

Because the variance of error scores within groups was not homogeneous, a logarithmic transformation was applied to the error scores before any analysis was undertaken. Since there were some zero error scores it was necessary to take the log of (X + 1) rather than log X. The logarithmic transformation served the additional purpose of reducing the influence of extreme scores upon the group means.

Equivalence of Experiments and Groups within Experiments. In order to test for equivalence of experiments and groups within experiments, a three-way analysis of variance was performed on the transformed error scores for the fourth training task. The requirement of homogeneity of variance for these scores was satisfied according to Bartlett's test

(1937). Table 3 summarizes the analysis.

TABLE 3

ANALYSIS OF VARIANCE FOR TRANSFORMED[a] ERROR SCORES FOR THE FOURTH TRAINING TASK

Source	df	ms	F
Blocks[b]	1	.014	.121
Rows	1	.696	5.645*
Columns	2	.066	.535
B x R	1	.119	.966
B x C	2	.142	1.152
R x C	2	.014	.120
R x C x B	2	.042	.338
Within	180	.123	
Total	191		

[a]Scores transformed by using log (X + 1).

[b]Blocks refers to experiments, row to type of training task, and columns to cue conditions.

*Significant at .05 level.

In Table 3 the only significant F value is associated with type of task. The mean of the transformed error scores for type W training tasks is significantly larger than the mean for type X tasks. This reflects the obvious difference in task difficulty. Since the analysis of variance failed to show any significant differences between experiments the results of the two experiments were pooled for all further analyses of error scores. Also since no significant differences were found for cue conditions (columns) it was concluded that the random assignments of Ss to groups had resulted in groups of similar ability.

Average Error Scores on the Criterion Task. Table 4 contains the means and standard deviations of the transformed error scores for

each group on the criterion task.

TABLE 4

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED[a] ERROR SCORES ON THE CRITERION TASK

Type of Training Task		Cue Condition S	O	N	Groups Pooled
W (N = 32 in each group)	M =	.651	.747	.589	.662
	S.D. =	.437	.311	.353	.373
X (N = 32 in each group)	M =	.841	.771	.843	.819
	S.D. =	.334	.440	.485	.421
Experimental Groups Pooled	M =	.746	.759	.716	.740
	S.D. =	.398	.378	.439	.404
None (Control group; N = 32)		--	--	--	1.267
		--	--	--	.440

[a]Scores transformed by using log (X + 1).

A three-way analysis of variance was used to test the predictions about group means. Again the logarithmic transformation which was used resulted in homogeneity of variance. Table 5 reports the analysis of variance.

The first prediction made concerning average group error scores on the criterion task involved the order of means for the three cue conditions. It was predicted that the order of error scores for cue groups (pooled) would be: smallest S, then N, largest O. These means are .746, .716, and .759. Differences between these means are not significant.

When average error scores for cue groups are compared independently

TABLE 5

ANALYSIS OF VARIANCE FOR TRANSFORMED[a] ERROR SCORES FOR THE CRITERION TASK

Source	df	ms	F
Blocks	1	.003	.016
Rows	1	1.170	7.245**
Columns	2	.031	.191
B x R	1	.122	.754
B x C	2	.028	.175
R x C	2	.224	1.383
R x C x B	2	.147	.911
Within	180	.162	
Total	191		

[a]Scores transformed by using log (X + 1).

[b]Blocks refers to experiments, row to type of training task, and columns to cue conditions.

**Significant at .01 level.

for each type of training task, there appears to be an interaction effect. Under the W training condition the O group made more errors than either other group while under the X training condition, the O group made fewer errors than either other group. However, the analysis of variance reported in Table 5 indicates that this interaction effect is not significant. The first prediction therefore was not confirmed. Although all of the experimental groups showed positive transfer when compared to the control group no net effect was found for cue conditions.

The second prediction, that the average error score for W groups would be less than that for X groups, was confirmed. The means for these groups were .662 and .819 respectively ($P < .01$). This is interpreted as showing the presence of a learning-to-learn effect.

In order to test for the presence of a warm-up effect a t test was used to compare the average transformed error score of the XN group with that of the control group. The means for these two groups, .843 and 1.267 differed significantly ($P < .005$).

Group Differences in Errors-Per-Trial. Within group error scores for each of the first 32 trials were computed. Graphic summaries by blocks of trials appear in Figures 4, 5, and 6 (see also Appendix E). The results are consistent with the analysis of variance (Table 5). In Figures 4 and 5 there is not a consistent difference in favor of any cue group. An interaction effect, possibly significant, appears. The WO group had some initial disadvantage on the criterion task, compared with WS and WN groups. The XO group, on the other hand, was not so handicapped compared with the XS or XN groups.

In Figure 6 the number of errors made for blocks of four trials by Ss in W groups is consistently less than the number of errors made by Ss in S groups. This again is evidence of a learning-to-learn transfer effect for Ss of W groups.

Differences in Cue Criterialities.

First-Trial Differences. In this study single-trial criterialities were proposed as a means of identifying sources for observed transfer effects. Primary interest therefore was centered on the initial trial of the criterion task. On the first trial of the criterion task each S made a response to a display containing four scaled cues. For each group of Ss a correlation coefficient was computed between cue values and responses (trial 1, Appendix E). With an N of 16 these criterialities

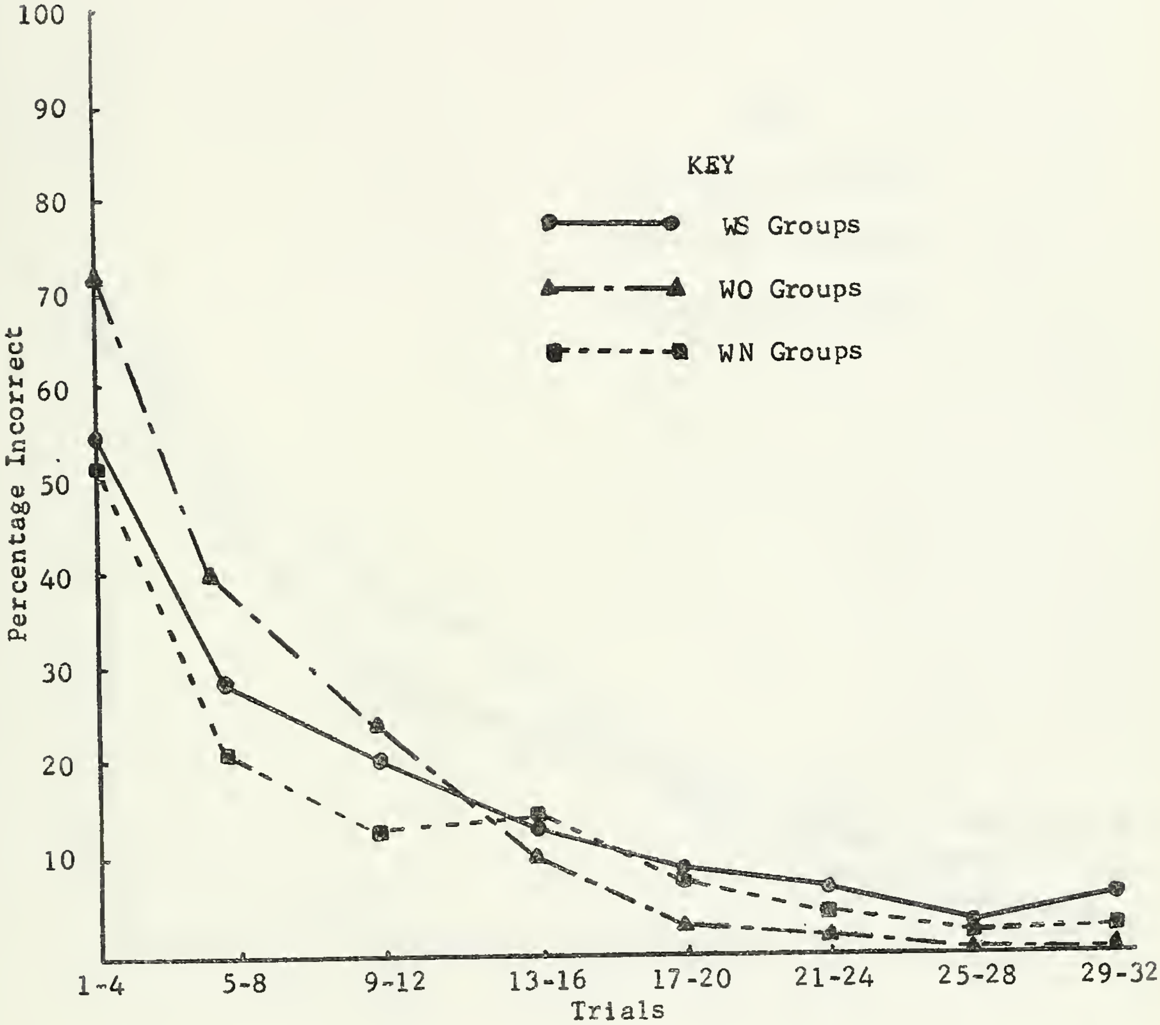


FIGURE 4
PERCENTAGE OF INCORRECT RESPONSES FOR BLOCKS OF FOUR TRIALS ON THE CRITERION TASK--WS, WO, AND WN GROUPS

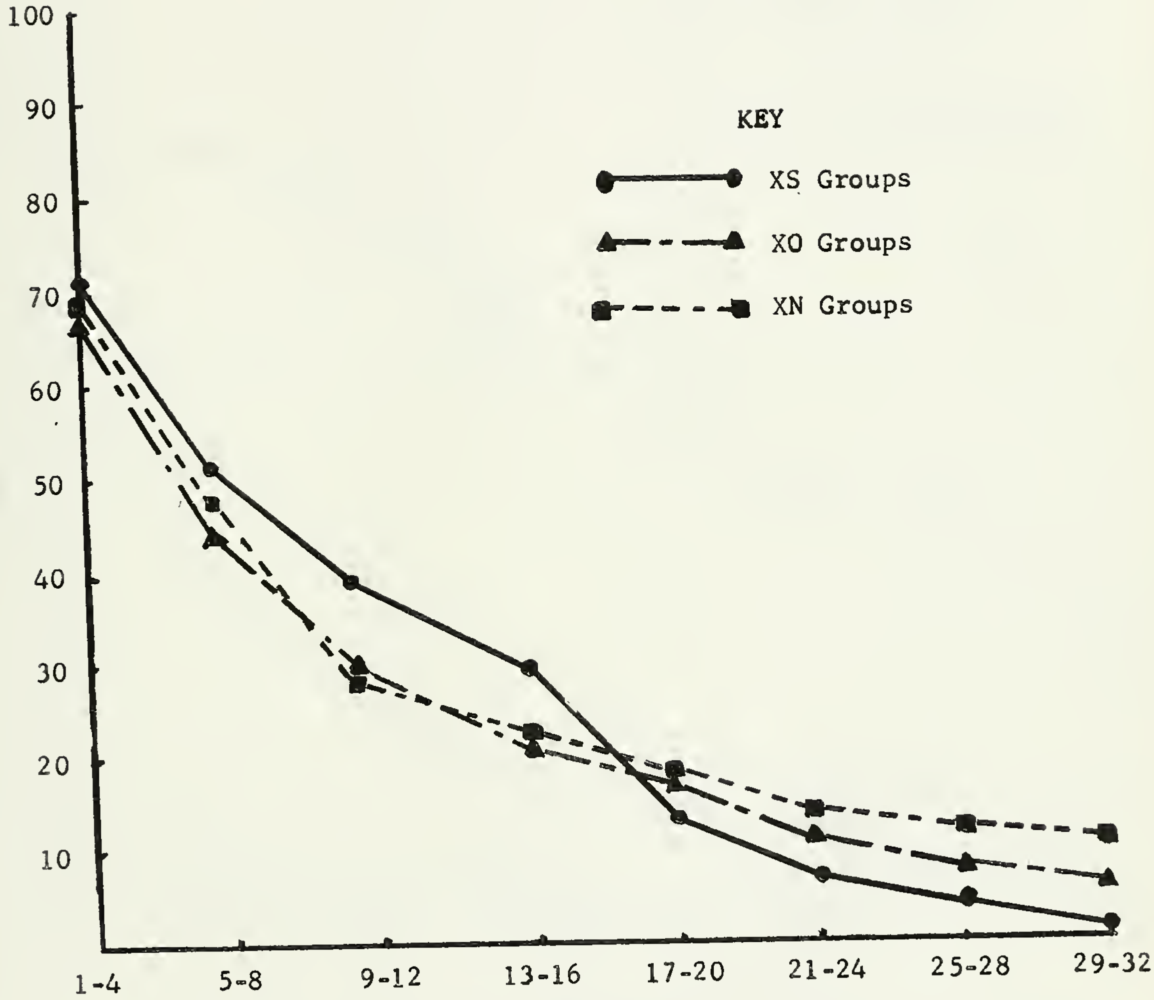


FIGURE 5
PERCENTAGE OF INCORRECT RESPONSES FOR BLOCKS OF FOUR TRIALS ON THE CRITERION TASK--XS, XO, AND XN GROUPS

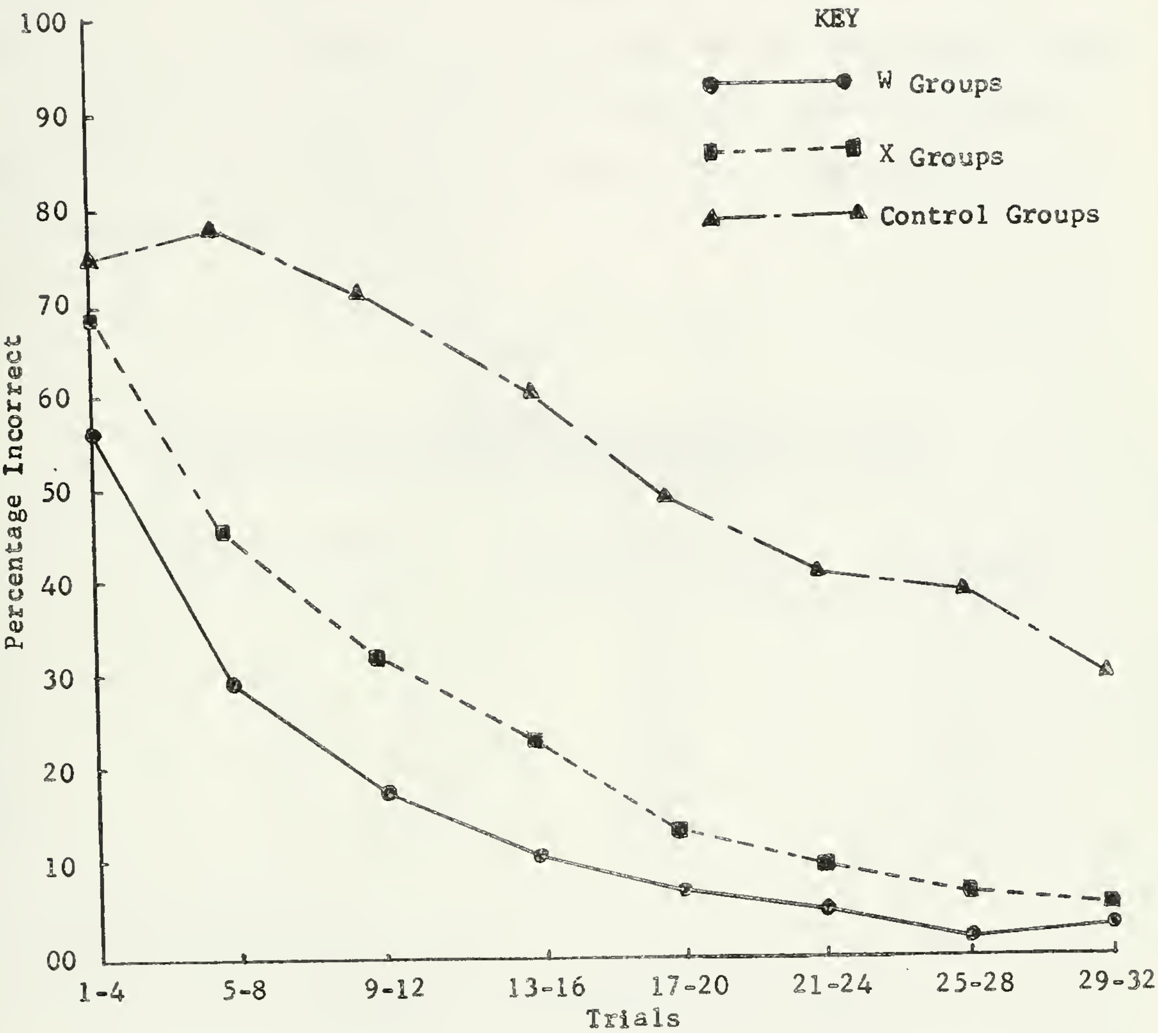


FIGURE 6
PERCENTAGE OF INCORRECT RESPONSES FOR BLOCKS OF FOUR TRIALS ON THE CRITERION TASK--W, X, AND CONTROL GROUPS
(Cue conditions Pooled)

are not very stable. Two steps were taken to increase N. First, data from the two experiments were pooled. Second, instead of computing the criteriality of each of the four cues separately, one criteriality was computed for the two relevant cues and another for the two irrelevant cues. This brought N for each criteriality to 64 (16 Ss in each of two experiments and two of each type of cue for each display). Table 6 contains the criterialities which resulted.

TABLE 6

CRITERIALITIES OF RELEVANT AND IRRELEVANT CUES ON THE FIRST TRIAL OF THE CRITERION TASK (N = 64[a])

Group	Cues Relevant on Criterion Task	Cues Irrelevant on Criterion Task	Difference[b]
WS	.27	.07	.20
WO	-.02	.44	.46
WN	.27	.11	.16
XS	.68	-.17	.85
XO	-.04	.29	.33
XN	.25	.14	.11
Control	.35	.04	.31

[a]Two entries per person.

[b]No exact significance test available (see page 33).

It was predicted that criterialities for cues which had been relevant on previous tasks would be larger than the criterialities for cues which had been irrelevant. For Table 6 this means that for WS and XS groups, relevant cues should have larger criterialities of the irrelevant cues; for WO and XO groups, criterialities of the irrelevant cues should be greater since relevant and irrelevant cues had been reversed. For all

four of these groups the results are in the predicted direction. For the control group and for the WN and XN groups no differences between the criterialities of the four cues on the first trial of the transfer task were expected.

There is not a suitable significance test for differences between criterialities because criterialities of the four cues are not independent. In order to make a test of significance possible the sum of relevant cues minus the sum of irrelevant cues was computed for each display. Correlating this sum with S's responses gave a sort of criteriality for relevant-minus-irrelevant cues. The values were: S groups, .38; O groups, -.40; and N groups, .14. With an N of 64 the values for S and O groups are significant at the .005 level. These results tend to confirm the predictions that were made.

The difference between criterialities of relevant and irrelevant cues for the control group (Table 6) is difficult to interpret. Nothing in the experimental design could account for this result. It must be assumed either that this is a chance effect or that some Ss looked at the first answer before responding. Some members of the control group may have inadvertently exposed the answer to the first display since they had not received previous training with the form of presentation used. An examination of errors made on the first few trials in the control group supports this explanation. Fewer errors were made on trial one than on any trial from two to nine. For other groups (See Appendix E) the greatest number of errors occurred on the first trial, except for group XN where there was one more error on trial two than on trial one.

Effects of Learning. The third prediction concerning cue criterialities was that the criterialities of relevant cues would approach the ideal values of .44 and .89 and the criterialities of irrelevant

cues would approach zero. Single-trial criterialities were computed for each group for every trial of the criterion task. Criterialities for the first 40 trials are reported in Appendix G. Criterialities for the remaining trials are not reported because beyond this point fluctuations were usually the result of changes in the responses of a single S. An examination of the criterialities in Appendix G reveals that for all groups the criterialities of relevant cues approached the ideal criterialities of .44 and .89 while criterialities of irrelevant cues approach zero.

The early effects of learning on the cue criterialities for the three cue conditions are shown in Figures 7, 8, and 9 (data in Appendix H). An analysis of additional trials would merely show a further increase in the criterialities of relevant cues and a decrease in the criterialities of irrelevant cues. These data were obtained by averaging the absolute values of the single-trial criterialities appearing in Appendix G. Two steps were taken in order to give these criterialities more stability. First, the two relevant and the two irrelevant cues were combined. Second, criterialities of cues were averaged for every two overlapping trials. Points on the graph in Figures 7, 8, and 9 thus represent the average of 16 criterialities (two experiments, two cues, two trials, and two training groups).

For all groups the average criterialities of relevant cues should increase and the criterialities of irrelevant cues should decrease. For S groups the criterialities of relevant cues should be higher initially than the criterialities of the irrelevant cues. For O groups the criterialities of relevant cues (previously irrelevant) should be initially lower than the criterialities of irrelevant cues (previously relevant).

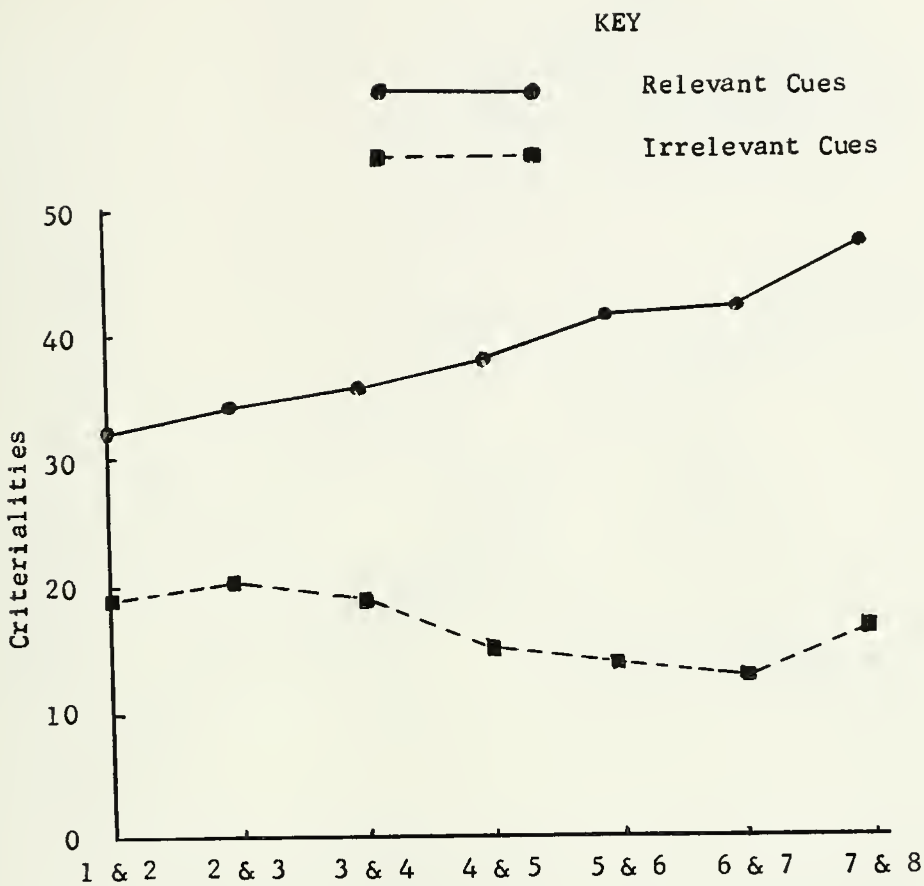


FIGURE 7

AVERAGE RELEVANT AND IRRELEVANT CUE CRITERIALITIES
FOR OVERLAPPING TRIALS ON THE CRITERION TASK--SAME CUE GROUPS

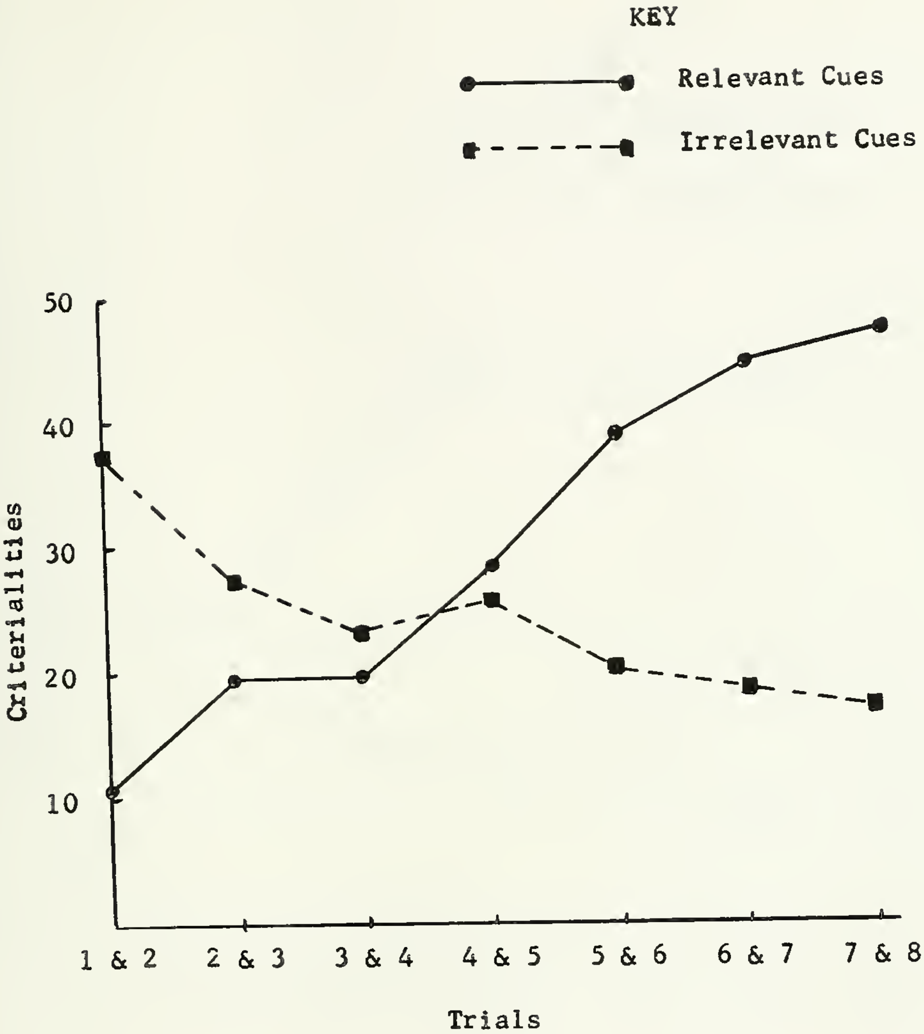


FIGURE 8

AVERAGE RELEVANT AND IRRELEVANT CUE CRITERIALITIES FOR OVERLAPPING TRIALS ON THE CRITERION TASK--OPPOSITE CUE GROUPS

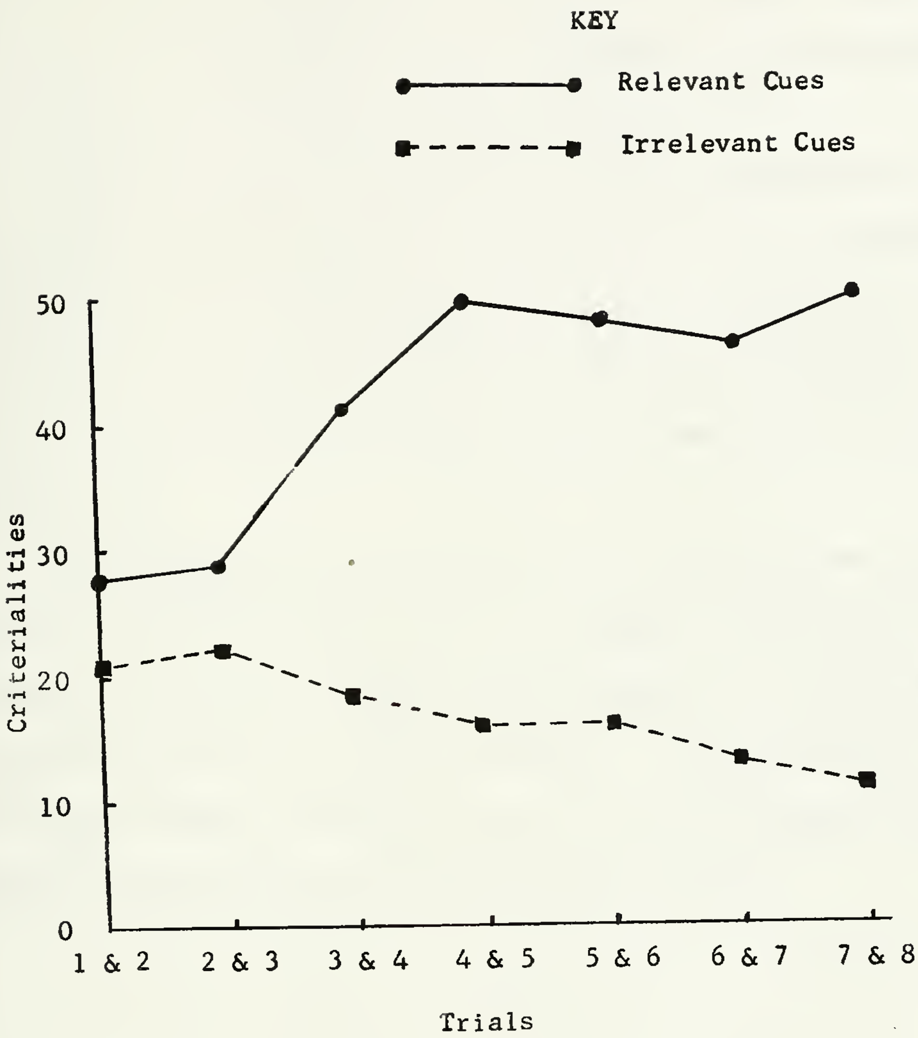


FIGURE 9

AVERAGE RELEVANT AND IRRELEVANT CUE CRITERIALITIES FOR OVERLAPPING TRIALS ON THE CRITERION TASK--NEW CUE GROUPS

For N groups the initial criterialities of relevant and irrelevant cues should be approximately the same. Only for the O groups should there be a crossover as learning takes place. The results shown in Figures 7, 8, and 9 conform very well with these expectations.

Figures 7, 8, and 9, together with the analysis of first trial criterialities, provide evidence of cue-repetition effects (facilitation in S, interference in O), consistent with the second hypothesis. These effects were not significant in Table 5, presumably because Ss had to know both which cues were relevant and how to weight these cues in order to avoid errors.

Single-subject criterialities were also computed for each S for the first block of 16 trials in order to compare the two kinds of criterialities (single-trial versus single-subject). The single-subject criterialities appear by group in Appendix I. For the purpose of this study these single-subject criterialities are not very meaningful because much learning occurred within the first block of 16 trials.

CHAPTER IV

DISCUSSION

The results of this experiment will be discussed in reference to the hypotheses tested and the predictions originally made.

Error Scores.

The first hypothesis was that through the use of the proper control groups it would be possible to identify and compare three kinds of transfer effects in a single study. The three kinds of transfer were the same as those described by Mandler (1962): an effect based on cue repetition, a learning-to-learn effect, and a warm-up effect. Related to the first is the second hypothesis, that for the effect associated with cue similarity both a negative and a positive transfer effect could be identified.

These two hypotheses resulted in the following predictions regarding group performances on the criterion task.

1. S groups would make fewer errors than N groups, whereas O groups would make more errors than N groups.
2. W groups would make fewer errors than X groups.

No prediction was made concerning the number of control group errors. This was, however, compared to the XN group (type X training tasks, new cues). Any difference in favor of the SN group would suggest a warm-up effect.

The predictions concerning the effects of cue repetition on error scores were not confirmed. Same cue groups, opposite cue groups and new

cue groups did not differ significantly. Several possibilities could account for this result.

The first possibility would be that the training which Ss received prior to the criterion task did not result in a bias in favor of certain cues for the criterion task. This possibility may be ruled out, in view of the criterialities of cues early in the criterion task.

Another possibility is that the instructions and the sample displays allowed Ss to guess correct weight before identifying relevant cues. It is also possible that the instructions and the feedback given after each trial provided enough information so that relevant cues were easily identified once correct weights were guessed. This explanation seems quite reasonable since an S who knew what weights to use and which cues were relevant would still have to guess which weights to apply to which cue. Another S knowing only what weights to use would have to choose two cues at random for the first trial. After receiving feedback in the form of the correct answer for the first trial it would be possible in many cases for both Ss to solve the problem. For example suppose Ss were shown a display containing cues with the following values 4, 3, 1, 2. If the correct answer to this display is 2.5 and S is quite sure that the weights are 1 and 1/2 only one solution is possible. The cue containing the 1 must be weighted 1/2 and the cue containing the 2 must be weighted 1.

One other possibility which may account for the failure to find significant differences between cue conditions is that the measure used as the rate of learning was not sensitive enough. Since time limits were not rigidly imposed for each trial, an error score of 4 for one S could represent 4 trials requiring 40 seconds and for another S the

same error score might represent 4 trials requiring four minutes.

Probably both of the factors which have been mentioned (amount of information given and lack of a strict control for time) contributed to the failure to find significant differences between cue conditions. This is a question which can be answered only by further studies.

The prediction regarding a learning-to-learn transfer effect was confirmed. The usual way to demonstrate learning to learn is to have Ss perform a series of related tasks of equal difficulty. A learning-to-learn curve can then be drawn by plotting the performance of S or a group of Ss for each task. In this study in which tasks were not of equal difficulty, a learning-to-learn curve would not be meaningful. However, if learning to learn is defined as better performance (compared with a control group) following performance of a series of related tasks, it should be possible to demonstrate this accumulated transfer effect at any point in the series. To avoid confounding learning-to-learn effects with warm-up effects it is necessary to provide this control group with a series of unrelated preliminary tasks. In this study W groups did better than X groups on the criterion task demonstrating the predicted accumulated learning-to-learn effect.

A warm-up effect would be expected to result in the XN group doing better on the criterion task than the control group. The difference found was substantial, indeed (see Table 4), larger than the learning-to-learn difference between the XN group and the WN group. This finding emphasizes the importance of controlling for warm-up when studying learning to learn.

It is important to control for learning-to-learn effects as well as for warm-up effects in the study of transfer associated with cue similarities.

A number of studies have compared reversal shifts and nonreversal shifts. In several of these studies an attempt has been made to determine the direction of transfer (D'Amato and Jagoda, 1960; Harrow and Friedman, 1958; Kendler and D'Amato, 1955; Kendler and Kendler, 1959; and Kelleher, 1959). Only Kelleher's control group received training on a task similar to the criterion task so that the effects of reversal and nonreversal shifts would not be confounded with learning-to-learn and warm-up effects. In Kelleher's study both reversal and nonreversal shifts produced negative transfer effects. In the other four studies the performance of Ss who had received training on one or more tasks was compared with that of Ss without previous training. Under these circumstances a positive transfer effect reported after a reversal shift might actually be a negative effect masked by positive learning-to-learn and warm-up effects. If, in our study, only the WO and control conditions had been employed, a positive transfer effect for a nonreversal shift would be reported rather than a finding of no significant difference.

Criterialities.

The third hypothesis was that training on a series of tasks in which the same cues were relevant would result in a bias in favor of the use of these cues on the criterion task and that this bias could be demonstrated by the use of single-trial criterialities. This hypothesis resulted in the following predictions:

1. For S and O groups single-trial criterialities for the first trial of the criterion task would be greater for cues relevant on preceding tasks than for cues irrelevant on preceding tasks.

2. For control groups and for N groups there would be no differences larger than chance expectancies among the single-trial criterialities of cues on the first trial of the criterion task.

3. For all groups single-trial criterialities for relevant cues would approach the ideal criterialities of .44 and .89. For nonrelevant cues the single-trial criterialities would approach zero.

The differences between criterialities of relevant and irrelevant cues on the first trial of the criterion task (Table 6) were consistent with the first prediction. No direct test for significant differences between these criterialities was possible. However an indirect test confirmed the prediction that Ss would use previously relevant cues in responding on the first trial of the criterion task ($P < .005$).

Since the second prediction was one of no significant difference this prediction could not be confirmed; it could only be disproved. For the control group the difference between the relevant and irrelevant cues on the first trial of the criterion task approached significance. This was interpreted as an artifact of the testing situation and not as evidence disconfirming the prediction.

The third prediction regarding cue criterialities was confirmed. For all groups the single-trial criterialities for cues tended to approach the ideal criterialities.

Single-trial criterialities provided useful information in this study, helping to explain the lack of significant differences among cue conditions. Since the criterialities indicated that Ss were biased in

their selection of cues on the criterion task some other explanation for the lack of significant differences in error scores was necessary. If the study were repeated, emphasis would be placed on getting a more sensitive measure of rate of learning and on reducing the amount of information given to Ss. No further effort would be made to make cues more distinctive or to increase the number of times the cues were used in training tasks.

Single-trial criterialities should be useful in other experiments besides those in which transfer effects are studied. Extended blocks of trials are not necessary in order to compute single-trial criterialities. This means that single-trial criterialities can be used for tasks requiring only a few trials to solution, if there is a suitable rotation of cue values over subjects.

Using single-trial criterialities it would also be possible to study the effect of a type of display within a sequence of displays. For example, in our W task, limits for correct answers were given at the start of each task. Therefore, Ss should be able to learn more from an initial display in which the correct answer is at either extreme than when it is in the middle of the range. When the correct answer is a maximum, only those cues containing maximum numbers can be relevant. Similarly, when K is a minimum only those cues containing minimum numbers can be relevant. To see whether Ss actually get more information from these displays a sequence of trials could be set up so that single-trial criterialities could be computed for the trial immediately following maximum or minimum K-displays and immediately following displays in which K values were of average size. These criterialities would be useful even when Ss had not yet discovered proper weights and so would give more information than counts of successes.

By using single-trial criterialities on the Azuma task it should be possible to study the rates at which Ss learn to respond correctly to the various classes of stimulus displays. In order to do this it would be necessary to develop a set of displays in which type of display is held constant for each trial within a group of Ss. Cronbach and Azuma (1961a) report that Ss divide the stimulus displays into four classes. A series of stimulus displays could be developed so that for every block of four trials the order of the four classes of displays is randomized or systematically varied. By using such a sequence of displays, single-trial criterialities could be used to compare the rates at which the four classes of displays are learned.

The title of one of the papers listed in the bibliography is "Can we tell what the learner is thinking from his behavior?" (Cronbach and Azuma, 1961b) The results of this study indicate that to a limited extent it is possible to identify what thoughts are most prominent within a group of learners by the use of single-trial criterialities. Even if completely accurate introspective reports were available it would be difficult to demonstrate the effect of a cue reversal any more clearly than the results which are shown in Figure 8.

Situations in which single-trial criterialities are obtainable are admittedly rather limited when compared with the wide range of tasks used in the investigation of human problem solving ability. Nevertheless the technique of this study does provide an objective method of identifying mediational processes during the solution of tasks similar to those of this study. As such it should be a useful research tool in the investigation of human problem solving ability.

Conclusions.

The study assessed independently three kinds of transfer effects; an effect associated with cue repetition, a learning-to-learn effect and a warm-up effect. In most studies these effects are confounded because of the lack of proper control groups. The importance of such controls was emphasized by our results. Using error scores as a measure of rate of learning, the greatest transfer effect was that attributed to warm-up. The next largest was a learning-to-learn effect. After both warm-up and learning-to-learn effects were eliminated no significant effect was found for cue repetition.

Our second major objective was to evaluate the usefulness of single-trial criterialities as indicators of the mediational processes of Ss. These single-trial criterialities proved to be useful in interpreting results. Moreover, they demonstrated the presence of cue-similarity effects not detectable in the error scores. Several other situations where single-trial criterialities would be useful were discussed.

CHAPTER V

SUMMARY

The first objective of this experimental study was to identify and compare three kinds of transfer effects: an effect associated with cue repetition, a learning-to-learn effect, and a warm-up effect. The second major objective was to evaluate the usefulness of cue-response criterialities in explaining transfer effects.

A factorial design was employed with three degrees of similarity between the relevant cues for the training tasks and those for the criterion task and two degrees of similarity between type of training task and criterion task.

Conditions of Cue Similarity. For one third of the experimental Ss relevant and irrelevant cues remained the same for all tasks; for another third relevant and irrelevant cues were reversed on the criterion task; and for the remaining third completely new cues were introduced during the criterion task.

Conditions of Task Similarity. For half of the experimental Ss training and criterion tasks were of the same type. For the other half training tasks and criterion tasks were quite different.

In addition to the six experimental groups necessary for the experimental design an additional group of Ss was a control group who performed only the criterion task.

The entire experiment was carried out twice--once using large group testing procedures and once testing groups of either 7 or 14 at a time.

The Ss for this experiment were undergraduate college students. For the first experiment in which large group testing procedures were used the Ss participated in the experiment as part of a course requirement either in introductory psychology or in educational psychology. For the second experiment all Ss volunteered.

The results of the study may be summarized as follows:

1. Ss of the three cue conditions did not differ significantly on the number of errors made during the completion of the criterion task.

2. A learning-to-learn effect was identified. Ss who received training on a series of training tasks similar to the criterion task completed the criterion task with fewer errors than Ss for whom training tasks were unlike the criterion task.

3. A warm-up effect was identified. Subjects who performed a series of four tasks quite different from the criterion task, using cues unlike those used on the criterion task, completed the criterion task with fewer errors than Ss in the control group.

4. Using the same two cues in the solution of a number of training tasks increased the use of these cues on the first trial of the criterion task. The criterialities (correlations between cues and responses) were higher on the first trial of the criterion task for cues that had previously been relevant than for cues that had been irrelevant.

BIBLIOGRAPHY

Adams, J. A. The second facit of forgetting: A review of warm-up decrement. *Psychol. Bull.* 1961, 58, 257-273.

Adams, J. A. Multiple versus single problem training in human problem solving. *J. exp. Psychol.*, 1954, 48, 15-18.

Adams, J. K. Laboratory studies of behavior without awareness. *Psychol. Bull.*, 1957, 54, 383-405.

Azuma, H. Comparison of a correlational with a probabilistic approach to concept learning. Unpublished doctoral dissertation, University of Illinois, 1960.

Bartlett, M. S. Some examples of statistical methods of research in agriculture and applied biology. *J. R. Statist. Soc. Suppl.*, 1937, 4, 137-170.

Bruce, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 133, 16, 343-361.

Bruner, J. S., Goodnow, J. J., and Austin, G. A. *A study of thinking.* New York: Wiley, 1956.

Brunswik, E. *Perception and the representative design of psychological experiments.* Berkeley, Calif.: University of California Press, 1956.

Brunswik, E. and Reiter, L. Eindruckscharaktere schematisierter gesichter. *J. Psychol.*, 1938, 142, 67-134.

Cox, J. W. Some experiments on formal training on the acquisition of skills. *Brit. J. Psychol.*, 1933, 24, 67-87.

Cronbach, L. J. *Educational Psychology*, N. Y.: Harcourt, 1963.

Cronbach, L. J. and Azuma, H. Can we tell what the learner is thinking from his behavior? Paper presented at the Conference on Research Methodology in Training, Washington University, St. Louis, Missouri, 1961. (a)

Cronbach, L. J. and Azuma, H. Performance on a concept-attainment task with scaled cues. Urbana, Illinois: University of Illinois, 1961. (Mimeo) (b)

D'Amato, M. R. and Jagoda, H. Transfer of brightness discrimination from bar pressing to a Y maze. *Psychol. Rep.*, 1960, 7, 83-90.

Davis, K. G., and Hess, H. F. The effectiveness of concepts at various levels of awareness. J. exp. Psychol., 1962, 63, 62-67.

Duncan, C. P. Description of learning how to learn in human subjects. Amer. J. Psychol., 1960, 73, 108-114.

Duncan, C. P. Transfer after training with single versus multiple tasks. J. exp. Psychol., 1958, 55, 63-72.

Eckstrand, G. A. and Wickens, D. D. Transfer of perceptual set. J. exp. Psychol., 1954, 47, 274-278.

Hamilton, C. E. The relationship between length of interval separating two learning tasks and performance on the second task. J. exp. Psychol., 1950, 40, 613-621.

Harlow, H. T. The formation of learning sets. Psychol. Rev., 1949, 56, 51-65.

Harlow, H. T., Harlow, M. K., Rueping, R. R., and Mason, W. A. Performance of infant rhesus monkeys on discrimination learning, delayed response, and discrimination learning set. J. comp. physiol. Psychol., 1960, 53, 113-121.

Harrow, M. and Friedman, G. B. Comparing reversal and nonreversal shifts in concept formation with partial reinforcement controlled. J. exp. Psychol., 1958, 55, 592-598.

Hull, C. L. Quantitative aspects of the evolution of concepts. Psychol. Monogr., 1920, 28, Whole No. 123.

Hunt, E. B. Concept Learning, N. Y.: Wiley, 1962.

Irion, A. L. The relationship of "set" to retention. Psychol. Rev., 1948, 55, 336-341.

Kelleher, R. T. Discrimination learning as a function of reversal and nonreversal shifts. J. exp. Psychol., 1956, 51, 379-384.

Kendler, H. H. and D'Amato, M. F. A comparison of reversal shifts and nonreversal shifts in human concept formation. J. exp. Psychol., 1955, 49, 165-174.

Kendler, H. H. and Kendler, T. S. Reversal and nonreversal shifts in kindergarten children. J. exp. Psychol., 1959, 58, 56-60.

Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. J. exp. Psychol., 1949, 39, 770-784.

Lawrence D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. J. exp. Psychol., 1950, 40, 175-188.

Mandler, G. The warm-up effect: Some further evidence on temporal and task factors. J. exp. Psychol., 1956, 55, 3-8.

Mandler, G. From association to structure. Psychol. Rev., 1962, 69, 415-427.

McHale, T. J. and Stolurow, L. M. Concept formation with metrically multi-valued cues and response categories. Urbana, Illinois: Training Research Laboratory, Tech. Rep. No. 3, ONR Proj., April, 1962.

Shepard, W. O. Learning set in preschool children. J. comp. physio. Psychol., 1957, 50, 15-17.

Smedslund, J. Multiple-probability learning: An inquiry into the origins of perception. Oslo, Norway: Akademisk, 1955.

Smoke, K. L. An objective study of concept formation. Psychol. Monogr., 1932, 42, Whole No. 191.

Stolurow, L. M., Solley, C. M., and Fisher, W. Problem solving in relation to prior probability learning. Urbana, Illinois: Memo. C-27 Task C, A. F. Proj., August 1955.

Thune, L. E. Warm-up effect as a function of level of practice in verbal learning. J. exp. Psychol., 1951, 42, 250-256.

Thune, L. E. The effect of different types of preliminary activities on subsequent learning of paired-associate material. J. exp. Psychol., 1950, 40, 423-438.

Thune, L. E. and Ericksen, S. C. Studies in abstraction learning: IV. The transfer effects of conceptual vs. rote instruction in a simulated classroom situation. Nashville: Vanderbilt Univ., 1960. (Mimeo)

Walk, R. D. Effect of discrimination reversal on human discrimination learning. J. exp. Psychol., 1952, 44, 410-419.

Walker, H. M., and Lev, J. Statistical Inference, New York: Holt, 1953.

Ward, L. B. Reminiscence and rote learning. Psychol. Monogr., 1937, 49, No. 220.

APPENDIX A

Order of Stimulus Cues for the Displays of Each Block

(16 Displays Per Block)

Display #	1st Figure	2nd Figure	3rd Figure	4th Figure
1.	rhombus[a]	circle	square	triangle
2.	square	circle	triangle	rhombus
3.	rhombus	square	circle	triangle
4.	circle	square	rhombus	triangle
5.	triangle	rhombus	circle	square
6.	square	rhombus	circle	triangle
7.	circle	triangle	rhombus	square
8.	triangle	square	circle	rhombus
9.	square	rhombus	triangle	circle
10.	circle	rhombus	triangle	square
11.	triangle	rhombus	square	circle
12.	rhombus	triangle	square	circle
13.	triangle	circle	square	rhombus
14.	circle	triangle	square	rhombus
15.	circle	rhombus	square	triangle
16.	rhombus	circle	triangle	square

[a]When the set of chemistry figures were used the following substitutions were made:

beaker - circle
funnel - square
flask - triangle
jar - rhombus

APPENDIX B

Description of Displays Used for Type X Tasks

Display #	1st Relevant Cue	2nd Relevant Cue	1st Irrelevant Cue	2nd Irrelevant Cue
1.	-[a]	-	X	X
2.	-	-	-	-
3.	X	-	-	X
4.	X	X	X	-
5.	X	-	X	-
6.	-	X	-	-
7.	-	-	-	X
8.	-	-	X	-
9.	X	X	X	X
10.	X	-	-	-
11.	X	-	X	X
12.	-	X	-	X
13.	X	X	-	X
14.	-	X	X	X
15.	-	X	X	-
16.	X	X	-	-
17.	-	X	X	-
18.	X	X	-	-
19.	-	-	-	-
20.	-	-	-	X
21.	-	X	-	-
22.	X	-	X	-
23.	X	X	X	X
24.	X	-	X	X
25.	-	X	-	X
26.	X	-	-	X
27.	X	-	-	-
28.	-	-	X	-
29.	X	X	-	X
30.	-	-	X	X
31.	-	X	X	X
32.	X	X	X	-
33.	-	X	X	-
34.	X	X	-	-
35.	-	-	-	-
36.	X	-	-	X
37.	X	-	-	-
38.	-	-	X	-
39.	X	X	X	-
40.	X	-	X	-

[a]X means figure contained an X; "-" means figure was empty.

APPENDIX B (Continued)

Display #	1st Relevant Cue	2nd Relevant Cue	1st Irrelevant Cue	2nd Irrelevant Cue
41.	-	-	-	X
42.	-	X	X	X
43.	X	X	-	X
44.	X	X	X	X
45.	-	X	-	X
46.	-	-	X	X
47.	-	X	-	-
48.	X	-	X	X
49.	X	-	X	-
50.	-	X	X	-
51.	X	X	-	-
52.	X	X	X	-
53.	X	X	X	X
54.	X	-	X	X
55.	-	-	X	X
56.	-	-	-	-
57.	X	-	-	X
58.	-	X	X	X
59.	-	-	-	X
60.	-	-	X	-
61.	-	X	-	X
62.	X	X	-	X
63.	-	X	-	-
64.	X	-	-	-
65.	-	-	X	X
66.	-	-	-	X
67.	X	X	X	-
68.	X	-	X	-
69.	-	-	-	-
70.	X	-	-	X
71.	-	X	X	X
72.	-	X	X	-
73.	-	-	X	-
74.	-	X	-	X
75.	X	X	-	X
76.	X	-	X	X
77.	X	X	-	-
78.	-	X	-	-
79.	X	-	-	-
80.	X	X	X	X

APPENDIX C

Instructions to Ss

The written instructions to the Ss varied for the two types of training tasks. The first set of instructions (labeled Appendix C-1) were used for W groups and the second set of instructions (labeled Appendix C-2) were used for X groups. A sheet of sample figures with their correct answers was included in each set of initial instructions. Following each set of initial instructions are the instructions which were given at the beginning of each of the subsequent tasks. New sample figures and their correct answers were also given at the beginning of each subsequent task. Since these sample sheets have been described in the main body of this report they do not appear here.

Control group Ss were given the instructions which appear in Appendix C-1.

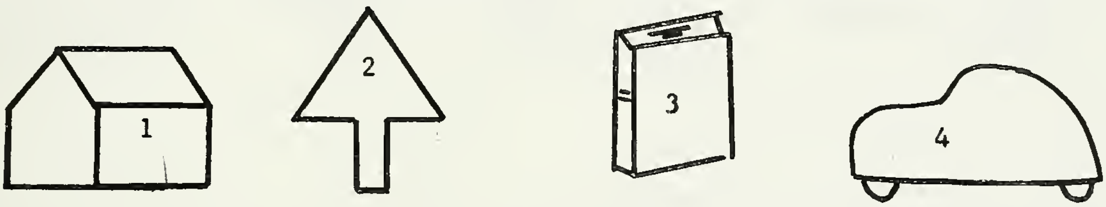
APPENDIX C-1
Instructions

(a) Please read these instructions carefully before beginning this problem. On all white sheets you are to look at frames in alphabetical order by beginning at the top frame on each page and proceeding to the bottom frame before turning to the next page.

Go to frame (b)

(b) This booklet contains a problem in which you are to learn to evaluate something called K-ness. You will see a series of figures each containing four shapes. Inside of each shape will be a number from one to four. Frame (c) contains a sample figure.

(c)



Sample Figure 1

(d) Your problem is to learn to estimate K for each figure. I will tell you two things about how K is determined. First, K is affected only by the numbers inside two of the four shapes. One of your tasks is therefore going to be to try to discover which two of the four shapes are relevant in determining K.

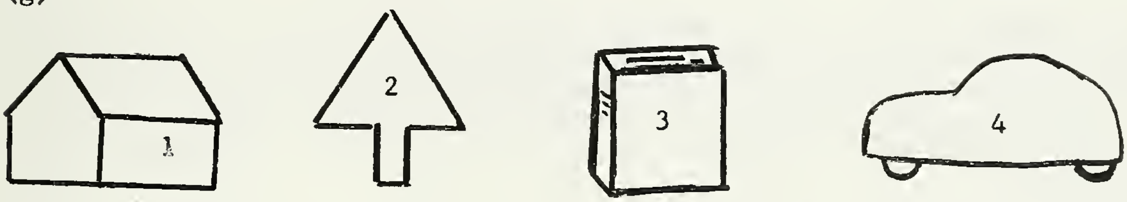
(e)

The second thing I will tell you about K is that K is always obtained by multiplying the numbers inside each of the relevant shapes by some constant and then adding. The numbers inside the two relevant shapes may or may not be multiplied by the same constant. Your second tasks is therefore to discover what constant to multiply each of the relevant shapes by.

Turn to frame (f) on page 2

(f) To make these instructions clear I am going to give you an example In this example the rule for determining K will be to multiply the number inside the house by four and the number inside the book by two before adding. Look at frames (g), (h), and (i).

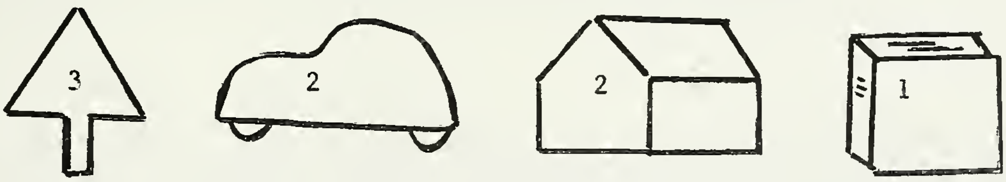
(g)



$K = 4 \times 1 + 2 \times 3$

$K = 4 + 6 = 10$

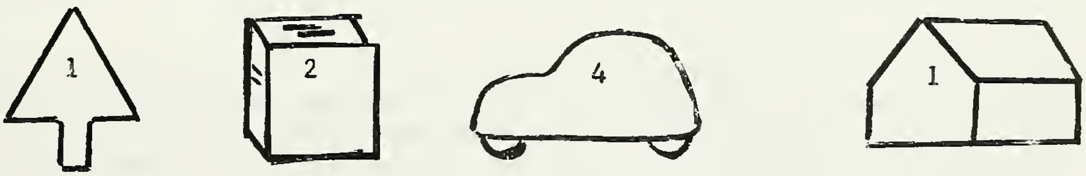
(h)



$K = 4 \times 2 + 2 \times 1$

$K = 8 + 2 = 10$

(i)



$K = 4 \times 1 + 2 \times 2$

$K = 4 + 4 = 8$

(j) In this example K was affected only by the numbers inside two figures; the house and the book. The number inside the house was always multiplied by the constant four and the number inside the book was always multiplied by the constant 2.

Go to Page 3 for frame (k)

(k) All of the pages following the next one will be blue pages containing problem figures. You are to look at only the top figure on each page before turning to the next page. When you get to the last blue page you are then to return to the first blue page and look at the second figure on each page. Do not be alarmed that the figures do not appear in correct numerical sequence. Rather they will appear in the same sequence as the numbers on your answer sheet.

(l) You are to look at each figure and then write in your estimate of K in the space opposite the appropriate figure number on the answer sheet. You are then to lift the blue tab on the left side of the booklet in order to see the correct K value. Your first estimates will necessarily be strictly guesses. Later you will learn to predict K correctly for each figure.

(m) When you get an answer correct put a large "C" by that answer. When you get 16 correct answers in a row you are to stop and write at the bottom of the answer sheet in the space provided the method which you are using to obtain K. You will then be through with this problem.

(n) You will be allowed to spend approximately twenty seconds on each figure. You should write in your answers near the beginning of this twenty second period so that you can see the correct answer while there is still some time remaining to study the figure. A bell will ring every twenty seconds indicating it is time to turn to the next figure. When you are making correct answers regularly you need not wait for the bell to ring before proceeding to the next figure.

(o) The next page contains four sample figures and their K values. These sample figures will give you some idea of the range of K for the figures in this problem. Spend about twenty seconds studying each of these sample figures. When you finish looking at the sample figures go to the first blue page of problem figures. Be sure that you answer only one frame on each page before going to the next page. Also be sure that you answer only one frame on each page before going to the next page. Also be sure that you write your answer on the line corresponding to the figure you are studying

SAMPLE FIGURES FOR TASK 1

SAMPLE FIG. 1[a]

K = 2

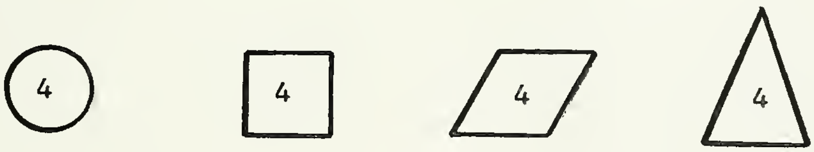
SAMPLE FIG. 2

K = 4

SAMPLE FIG. 3

K = 6

SAMPLE FIG. 4



K = 8

For this TASK, K will not be greater than 8 or less than 2.

[a]For new cue groups the appropriate chemistry symbols were used as figures.

Instructions

In this new problem K is determined a little differently. Again only two shapes are relevant and K is obtained by multiplying the numbers inside each of the two relevant shapes by some constant. Follow the same procedure as for the previous task.

APPENDIX C-2
Instructions

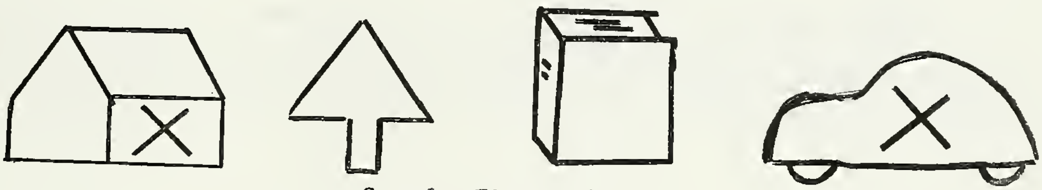
(a) Please read these instructions carefully before beginning this problem. On all white sheets you are to look at frames in alphabetical order by beginning at the top frame on each page and proceeding to the bottom frame before turning to the next page.

Go to frame (b)

(b)

This booklet contains a problem in which you are to learn to evaluate something called "K". You will see a series of figures each containing four shapes. Each shape will either contain an "X" or it will be empty. Frame (c) contains a sample figure.

(c)



Sample Figure 1

(d)

Your problem is to learn to judge whether each figure is a K figure or not. I will give you a hint. Although there are four shapes in each figure, only two shapes influence whether a figure is a K figure or not. For example, suppose only the house and the book are relevant in determining whether a figure is a K figure or not. This means that you could entirely disregard the car and the tree. One of your tasks for this problem will therefore be to discover which two shapes are relevant.

(e)

Your second task will be to discover what it is about these two relevant shapes which determines whether or not the figure is a K figure. Order of shapes will be irrelevant so that the presence or absence of X's in the two relevant shapes must be the basis for classifying figures. It may be that both relevant shapes must be empty, or that both must contain X's, or that one must be empty and the other contain an X.

(f)

All of the pages following the next one will be blue pages containing problem figures. You are to look at only the top figure on each page before turning to the next page. When you get to the last blue page you are then to return to the first blue page and look at the second figure on each page. Do not be alarmed that the figures do not appear in correct numerical sequence. Rather they will appear in the same sequence as the numbers on your answer sheet.

(g)

You are to look at each figure and then guess whether or not the figure is a K figure. If you think the figure is a K figure write a K on the space opposite the appropriate figure number on the answer sheet. If you do not believe the figure is a K figure make an O on your answer sheet. Your first answers will be strictly guesses. Later you will learn to classify each figure correctly. In order to find the correct answer for each figure lift the blue tab on the left side of the booklet.

(h)

When you get an answer correct put a large "C" by that answer. When you get 16 correct answers in a row you are to stop and write at the bottom of the answer sheet in the space provided a description of how a K card is defined. You will then be through with this problem.

(i)

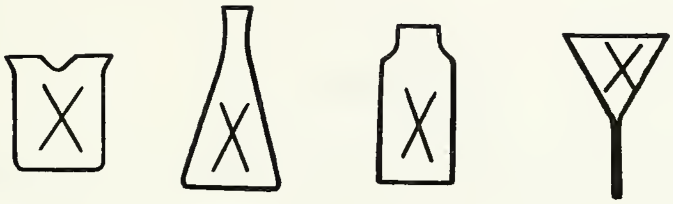
You will be allowed to spend approximately twenty seconds on each figure. You should write in your answers near the beginning of this twenty second period so that you can see the correct answer while there is still some time remaining to study the figure. A bell will ring every twenty seconds indicating it is time to turn to the next figure. When you are making correct answers regularly you need not wait for the bell to ring before proceeding to the next figure.

(j)

The next page contains two sample figures and their correct answers. These sample figures will give you some idea of how a K card is defined in this problem. Spend about twenty seconds studying each of the sample figures. When you finish looking at the sample figures go to the first blue page of problem figures. Be sure that you answer only one frame on each page before going to the next page. Also be sure that you write your answer on the line corresponding to the figure you are studying.

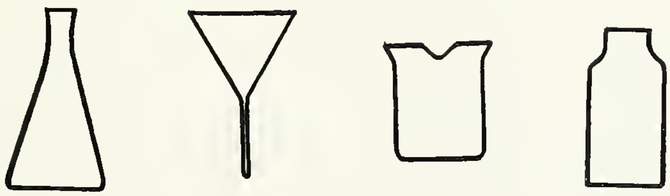
SAMPLE FIGURES FOR TASK 1

SAMPLE FIGURE 1[a]



Answer K

SAMPLE FIGURE 2



Answer O

[a]For same cue groups and opposite cue groups the appropriate geometric figures were used.

Instructions

In this new problem a K figure is defined a little differently. Again only two shapes are relevant and it is the presence or absence of X's in these relevant shapes that determines whether a figure is a K figure. Follow the same procedure as for the previous task.

Instructions

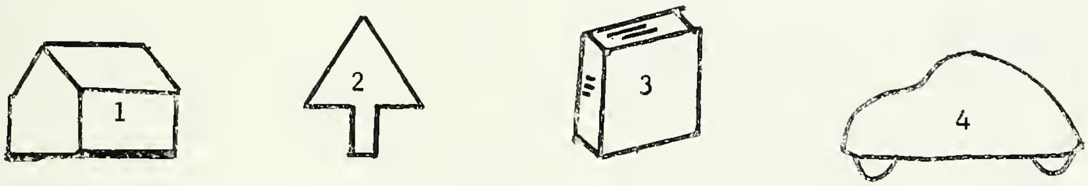
(a)

Now you are going to be asked to solve a task which is of a different type from those which you have been solving. You will, however, discover some similarities.

(b)

In this problem you are going to learn to evaluate something called K-ness. You will se a series of displays each containing four shapes. Inside each shape will be a number from one to four. Frame (c) contains a sample figure.

(c)



Sample Figure 1

(d)

Your problem is to learn to estimate K for each figure. I will tell you two things about how K is determined. First, K is affected only by the numbers inside two of the four shapes. One of your tasks is therefore going to be to try to discover which two of the four shapes are relevant in determining K.

(e)

The second thing I will tell you about K is that K is always obtained by multiplying the numbers inside each of the relevant shapes by some constant and then adding. The numbers inside the two relevant shapes may or may not be multiplied by the same constant. Your second task is therefore to discover what constant to multiply each of the relevant shapes by.

Turn to frame (f) on page 2

(f)

To make these instructions clear I am going to give you an example. In this example the rule for determining K will be to multiply the number inside the house by four and the number inside the book by two before adding. Look at frames (g), (h), and (i).

(g)

$K = 4 \times 1 + 2 \times 3$

$K = 4 + 6 = 10$

(h)

$K = 4 \times 2 + 2 \times 1$

$K = 8 + 2 = 10$

(i)

$K = 4 \times 1 + 2 \times 2$

$K = 4 + 4 = 8$

(j)

In this example K was affected only by the numbers inside two figures: the house and the book. The number inside the house was always multiplied by the constant four and the number inside the book was always multiplied by the constant 2.

Go to Page 3 for frame (k)

(k)

You are to follow the same general procedure for this task as for earlier tasks. For this task your answers will be numbers rather than K's and O's. Again you will have twenty seconds for each display. Answers will again be found under the blue tabs at the left side of the booklet. Mark correct answers with a "C" then when you get 16 correct answers in a row write the formula for determining K in the space provided at the bottom of the answer sheet.

(l)

The next page contains four sample figures and their K values. These sample figures will give you some idea of the range of K for the figures in this problem. Spend about twenty seconds studying each of these sample figures. When you finish looking at the sample figures go to the first blue page of problem figures. Be sure that you answer only one frame on each page before going to the next page. Also be sure that you write your answer on the line corresponding to the figure you are studying.

SAMPLE FIGURES FOR TASK 5

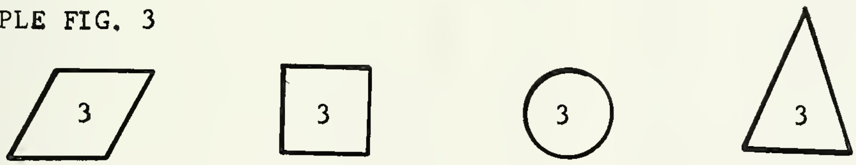
SAMPLE FIG. 1

K = 1½

SAMPLE FIG. 2

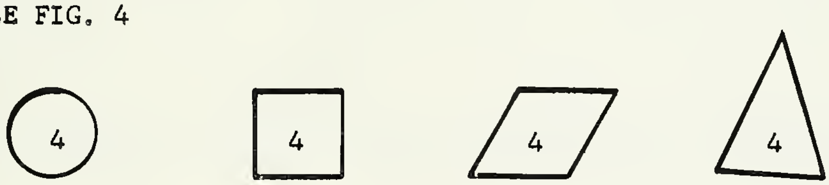
K = 3

SAMPLE FIG. 3



K = 4½

SAMPLE FIG. 4



K = 6

For this TASK, K will not be greater than 6

or less than 1½.

APPENDIX D

Error Scores for Each S by Group

Group WS

S#	Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter-ion task	Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter-ion
1	13	16	15	5	49	10	10	23	5	4	42	12
2	18	59	57	12	146	5	13	1	1	2	17	2
3	7	19	14	7	47	0	1	11	1	1	14	1
4	3	5	11	1	20	16*	4	4	2	2	12	4
5	2	9	9	3	23	3*	20	6	7	24	57	21
6	2	1	5	6	14	3	9	2	15	9	35	2
7	4	24	50	6	84	12	3	26	21	11	61	2
8	39	0	8	5	52	0*	2	8	17	1	28	41*
9	1	6	13	4	24	1	42	6	32	2	82	2
10	6	0	2	0	8	0	6	21	2	5	34	11
11	25	20	3	1	49	2	2	2	1	1	6	3
12	11	56	61	14	142	5	7	5	2	1	15	0
13	6	10	31	7	54	27	5	9	1	3	18	4
14	6	1	13	10	30	4	2	3	2	1	8	0
15	16	15	1	5	37	2	65	10	2	26	103	1
16	1	1	70	1	73	5	4	2	4	10	20	6
Med.	6.17	9.50	13.00	5.17	48.00	3.50	5.50	6.00	2.30	2.50	24.00	2.50
Mean	10.00	15.12	22.68	5.44	53.25	5.94	12.19	8.69	7.19	6.44	34.50	7.00
S.D.	10.35	18.29	23.3	3.99	40.96	7.09	17.37	7.91	9.19	7.99	28.12	10.64

*The S originally scheduled to be a member of this group had to be eliminated either because he failed to follow directions or because he failed to complete the training tasks. Data reported is for a second S.

APPENDIX D (Continued)

Group WO

S#	Experiment I Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criterion Task	Experiment II Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criterion
1	50	13	63	8	134	6	16	7	2	3	28	8
2	3	46	53	2	104	0	3	3	3	3	12	2
3	6	4	12	16	38	11	5	2	7	2	16	4
4	18	0	3	3	24	11	17	4	3	7	31	9
5	12	2	1	13	28	5	11	3	34	13	61	9
6	10	13	2	72	97	5	7	18	47	5	77	7
7	4	2	1	22	29	7	2	15	20	8	45	8*
8	20	6	3	2	31	7*	18	0	0	17	35	20
9	8	7	7	4	26	2*	3	3	2	4	12	3
10	7	17	6	2	32	1	9	6	48	2	65	11
11	48	17	5	1	71	8	12	13	9	2	36	2
12	9	3	16	9	37	0*	7	3	0	1	11	4
13	4	1	1	2	8	2	4	7	2	7	20	4
14	0	1	0	1	2	3	5	1	7	2	15	10
15	7	6	1	5	19	2	13	0	14	17	44	5
16	6	1	8	2	17	11	14	7	7	20	48	2
Med.	7.5	5.0	4.0	3.5	30.0	4.00	8.0	3.5	6.83	4.50	33.0	6.00
Mean	13.25	8.69	11.38	10.25	43.56	5.06	9.12	5.75	12.81	7.06	34.75	6.75
S.D.	14.88	11.49	18.81	17.57	37.64	3.86	5.37	5.35	16.12	6.25	20.62	4.67

APPENDIX D (Continued)

Group WN

S#	Experiment I Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criterion Task	Experiment II Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criterion
1	2	3	2	5	12	4	47	30	18	0	95	1
2	4	0	8	1	13	3	5	12	7	5	29	2
3	27	19	7	1	54	7	2	6	68	2	78	3
4	34	3	71	11	119	1	1	26	14	19	60	2
5	40	3	17	8	68	10	8	3	6	4	21	0
6	3	17	3	7	30	1	10	13	0	10	33	5
7	1	1	1	5	8	1	10	2	24	8	44	1
8	3	1	0	1	5	16	7	1	4	46	58	17
9	1	1	5	7	14	3	10	12	15	15	52	2
10	26	6	6	5	43	13	4	4	4	4	16	4
11	15	4	8	0	27	1	52	4	19	9	84	0
12	1	2	11	1	15	2	1	3	3	2	9	4
13	37	0	0	2	39	3	47	21	0	4	72	1
14	3	2	3	1	9	1	1	2	9	6	18	2
15	9	15	2	9	35	1	47	4	3	6	60	28
16	11	6	2	3	22	3	4	1	0	4	9	6
Med.	6.5	2.5	4.0	4.0	24.5	2.75	9.83	3.83	6.5	5.5	59.0	2.25
Mean	13.56	5.19	9.12	4.19	32.06	4.38	16.00	9.00	12.12	9.00	46.12	4.87
S.D.	14.29	6.16	17.1	3.44	29.11	4.68	19.51	9.27	16.67	11.01	27.84	7.37

APPENDIX D (Continued)

Group XS

	Experiment I						Experiment II					
S#	Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter-ion Task	Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter-ion
1	12	8	17	3	40	14	22	1	2	7	32	6
2	33	21	24	5	83	24	3	4	1	3	11	3
3	7	7	9	8	31	14	2	6	68	2	78	3
4	8	3	7	3	21	17	4	2	4	2	12	2
5	23	33	6	10	72	4	15	1	13	16	45	12
6	6	18	8	2	34	18	26	4	1	6	37	2
7	3	2	2	4	11	1*	4	5	1	3	13	4
8	1	1	0	3	5	3*	10	4	8	2	24	7
9	3	8	1	1	13	3	21	5	6	4	36	15
10	59	0	17	16	92	16*	30	4	6	0	40	6
11	19	5	7	2	33	4	5	3	1	0	9	1
12	29	16	8	3	56	25*	39	3	14	2	58	5
13	17	3	5	6	31	3	12	2	1	0	15	10
14	5	4	1	0	10	2	23	15	11	0	49	11
15	25	2	23	4	54	6	18	11	2	5	36	15
16	20	9	2	2	33	4	44	12	10	15	81	1
Med.	14.5	6.0	7.0	2.75	33.0	5.00	16.5	6.17	4.0	2.5	36.0	5.50
Mean	16.88	8.75	8.56	4.50	38.69	9.88	17.38	5.12	9.31	4.19	36.00	6.43
S.D.	15.05	8.98	7.69	3.99	26.27	8.24	12.96	4.06	16.30	4.90	22.51	4.77

APPENDIX D (Continued)

Group XO

	Experiment I						Experiment II					
S#	Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criterion Task	Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criterion
1	10	10	24	4	48	22	5	2	2	4	13	2
2	16	15	2	4	37	7	15	9	1	1	26	2
3	11	16	21	12	60	3*	4	2	2	4	12	3
4	30	18	12	5	65	1*	16	19	9	5	49	17
5	24	2	4	2	32	1	15	0	14	5	34	6
6	3	28	3	3	37	5	4	14	4	20	42	1
7	33	5	3	4	45	4	11	2	7	6	26	23
8	43	8	12	2	65	14	11	3	1	0	15	8
9	3	6	4	5	18	10	8	2	1	5	16	1
10	1	28	37	8	74	9	8	3	0	10	21	20
11	6	8	18	0	32	0	25	3	1	1	30	4
12	22	7	19	3	51	7	6	3	6	8	23	1
13	1	30	12	1	44	2	8	3	1	2	14	1
14	1	34	7	1	43	1	3	4	6	1	14	5
15	25	27	26	7	85	74	2	4	3	0	9	3
16	1	11	7	6	25	38*	13	36	2	5	56	5
Med.	10.5	13.0	12.17	4.17	43.5	6.00	7.83	2.90	2.17	4.5	22.0	3.50
Mean	14.38	15.81	13.19	4.19	47.56	12.38	9.62	6.81	3.75	4.81	25.00	6.37
S.D.	13.54	10.39	9.68	3.04	18.27	19.12	6.10	9.25	3.78	4.94	14.01	7.14

APPENDIX D (Continued)

Group XN

S#	Experiment I Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter- ion Task	Experiment II Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter- ion
1	14	5	1	1	21	3	18	1	2	7	28	5
2	5	7	22	12	46	6	11	10	12	2	35	1
3	26	9	11	3	49	3	12	2	9	7	30	4
4	21	6	27	1	55	1	1	6	4	2	13	1
5	19	19	21	12	71	19	19	3	1	5	28	6
6	14	7	13	1	35	4	52	5	2	2	61	9
7	44	7	24	6	81	19	9	5	2	6	22	0
8	16	21	12	9	58	7	8	2	2	2	14	3
9	16	22	3	6	47	2	19	15	7	6	47	5
10	11	8	0	2	21	15*	9	3	3	0	15	11
11	26	6	4	10	46	2	24	5	2	3	34	36
12	7	11	14	2	34	0	11	4	5	1	21	3
13	7	5	33	2	47	9	2	20	6	7	35	7
14	12	6	1	0	19	10	7	1	6	4	18	9
15	32	31	5	6	74	72	23	0	13	3	39	73
16	2	1	0	0	3	1	14	29	25	4	72	53
Med.	15.0	6.83	12.0	2.5	46.5	5.00	11.5	4.5	4.5	3.5	29.0	5.50
Mean	17.00	10.69	11.94	4.56	44.19	10.81	14.94	6.94	6.31	3.81	32.00	14.12
S.D.	10.91	8.13	10.71	4.23	21.46	17.46	11.96	7.96	6.18	2.29	16.66	21.13

APPENDIX D (Continued)

Control Groups

S#	Experiment I Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter-ion task	Experiment II Task 1	Task 2	Task 3	Task 4	Total 1 - 4	Criter-ion Task
1						15						43
2						43						20
3						7						12
4						55						15
5						18						17
6						74						0
7						40						20
8			No			20			No			63
9						67						8
10			Training			14			Training			6
11						5						9
12			Tasks			68			Tasks			12
13						4						36
14						10						4
15						13						29
16						17						73
Med.						17.50						16.00
Mean						29.38						22.94
S.D.						24.57						21.08

APPENDIX E

Total Number of Errors by Trial for the First 32 Trials of the Criterion Task

	Groups						
Trial	WS	WO	WN	XS	XO	XN	Control
1.	21[a]	28	20	24	26	25	22
2.	21	27	18	25	22	23	26
3.	12	21	15	23	18	23	24
4.	15	17	12	18	20	17	24
5.	12	14	9	20	19	17	26
6.	11	12	8	17	13	15	23
7.	6	12	6	12	14	16	25
8.	8	13	3	17	13	13	26
9.	6	11	3	13	12	9	21
10.	7	10	4	15	10	10	24
11.	7	6	4	15	9	10	22
12.	6	4	4	9	8	7	24
13.	6	5	5	9	8	7	20
14.	4	3	4	9	9	7	20
15.	3	1	4	10	8	7	18
16.	3	1	2	8	4	8	19
17.	3	1	3	4	5	5	15
18.	3	1	3	4	5	4	17
19.	3	1	2	2	5	5	17
20.	3	1	2	4	4	6	14
21.	4	1	3	3	6	5	14
22.	1	1	2	1	2	5	14
23.	1	0	2	2	4	5	13
24.	3	0	1	2	3	3	12
25.	2	0	1	1	2	4	13
26.	1	0	0	1	3	4	12
27.	0	0	0	1	2	4	13
28.	0	0	1	1	1	4	13
29.	2	0	1	0	2	4	8
30.	2	0	1	0	2	3	11
31.	2	0	1	0	1	3	10
32.	1	0	1	0	2	4	9

[a]Since there were 32 _Ss_ in each group when the two experiments were combined, the maximum number of errors possible for any trial would be 32.

APPENDIX F

Percentage of Incorrect Responses for Blocks of Four Trials on the Criterion Task (Figures 4, 5, and 6)

				Groups					
Trials	W	X	WX	WO	WN	XS	XO	XN	Control
1- 4	.56	.69	.54	.73	.51	.70	.67	.69	.75
5- 8	.29	.46	.29	.40	.20	.52	.46	.48	.78
9-12	.19	.33	.20	.24	.12	.39	3.0	.28	.71
13-16	.11	.25	.12	.08	.12	.28	.23	.23	.60
17-20	.07	.14	.09	.03	.08	.11	.15	.16	.49
21-24	.05	.11	.07	.02	.06	.06	.12	.14	.41
25-28	.01	.07	.02	.00	.02	.03	.06	.12	.40
29-32	.03	.06	.05	.00	.03	.00	.05	.11	.30

APPENDIX G

Single-Trial Cue Criterialities

Group WS

Trials	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
1	.83	.17	-.07	-.10	.02	.18	.00	.51
2	.19	.31	-.13	.29	.67	.34	.20	-.02
3	.19	.40	.46	.11	.57	.74	-.07	.07
4	.36	.74	.08	-.02	.41	.41	.19	-.30
5	.43	.75	-.14	-.09	.27	.40	.05	.44
6	.50	.36	-.04	.16	.19	.82	.12	.29
7	.30	.82	-.18	.11	.52	.76	-.02	-.16
8	.13	.91	-.09	.07	.36	.78	-.25	-.02
9	.32	.86	.02	-.22	.65	.71	.06	-.04
10	.37	.80	-.11	.15	.47	.67	-.03	.08
11	.60	.73	.03	-.10	.47	.61	-.24	.26
12	.45	.80	.00	-.13	.09	.55	.45	-.07
13	.06	.68	.48	.18	.50	.82	.05	-.08
14	.40	.89	-.01	-.05	.46	.84	.04	-.04
15	.62	.68	.07	.09	.44	.89	-.03	-.01
16	.48	.85	.10	-.05	.42	.77	-.18	.01
17	.47	.86	.00	.16	.34	.90	-.04	.31
18	.39	.90	-.16	.15	.40	.77	-.19	.31
19	.26	.87	-.24	.14	.37	.84	.02	.28
20	.55	.80	-.12	.18	.47	.86	-.11	.28
21	.41	.86	-.05	.31	.50	.76	.06	.07
22	.45	.89	-.08	.21	.42	.90	-.05	.22
23	.55	.80	.03	.32	.45	.89	-.08	.21
24	.47	.88	-.07	.22	.58	.58	.11	.29
25	.52	.82	-.12	.10	.50	.85	-.10	.20
26	.45	.89	-.10	.23	.45	.89	-.08	.21
27	.45	.89	-.08	.21	.45	.89	-.08	.21
28	.45	.89	-.08	.21	.45	.89	-.08	.21
29	.46	.88	-.05	.18	.46	.89	-.07	.22
30	.48	.85	-.16	.20	.48	.87	-.10	.18
31	.41	.89	-.06	.14	.45	.89	-.09	.22[a]
32	.50	.82	-.05	.25	.45	.89	-.08	.21
33	.45	.89	.02	-.03	.45	.89	.02	-.03
34	.45	.89	.02	.03	.46	.89	.02	.04
35	.45	.89	.02	-.01	.45	.89	.02	.00
36	.45	.89	.02	-.13	.45	.89	.02	-.13
37	.45	.89	.02	-.07	.45	.89	.03	-.10
38	.45	.89	.02	-.10	.50	.85	.07	-.04
39	.45	.89	.02	-.04	.45	.89	.04	-.04
40	.45	.89	.02	-.12	.48	.87	.03	-.16

[a]Correlations for nonrelevant cues are above .10 for the second block of trials due to an error in the cue values listed by McHale & Stolurow (1962).

APPENDIX G (Continued)

Group WO

Trials	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
1	.01	.03	.68	.29	-.05	.03	.54	.22
2	-.01	.17	.07	.78	.18	-.18	.46	.32
3	.11	-.22	.39	.44	-.04	.27	.16	.00
4	-.03	.23	.35	-.28	-.01	.26	.58	.28
5	.14	.62	-.27	-.29	-.06	.55	.34	.52
6	.22	.56	.16	.41	.22	.60	.11	.32
7	.26	.66	.15	.10	.46	.72	.10	-.10
8	.44	.75	.14	.22	.42	.77	-.06	.15
9	.12	.54	.20	-.12	.41	.51	.23	.27
10	.05	.71	.14	.28	.31	.89	.04	.09
11	.42	.88	-.09	-.02	.52	.80	-.01	-.06
12	.36	.92	.02	.02	.36	.57	.08	-.08
13	.43	.80	-.18	.02	.22	.78	.24	.20
14.	.45	.89	.00	-.04	.48	.85	-.02	.02
15	.45	.89	.00	-.04	.47	.88	-.03	-.06
16	.45	.89	.00	-.04	.35	.86	.05	-.10
17	.45	.89	-.08	.21	.44	.89	-.07	.20
18	.45	.89	-.08	.21	.47	.84	.02	.10
19	.45	.89	-.08	.21	.50	.82	-.19	.19
20	.45	.89	-.08	.21	.47	.84	-.11	.31
21	.45	.89	-.08	.21	.48	.85	-.06	.24
22	.45	.89	-.08	.21	.41	.89	-.12	.24
23	.45	.89	-.08	.21	.45	.89	-.08	.21
24	.45	.89	-.08	.21	.45	.89	-.08	.21
25	.45	.89	-.08	.21	.45	.89	-.08	.21
26	.45	.89	-.08	.21	.45	.89	-.08	.21
27	.45	.89	-.08	.21	.45	.89	-.08	.21
28	.45	.89	-.08	.21	.45	.89	-.08	.21
29	.45	.89	-.08	.21	.45	.89	-.08	.21
30	.45	.89	-.08	.21	.45	.89	-.08	.21
31	.45	.89	-.08	.21	.45	.89	-.08	.21
32	.45	.89	-.08	.21	.45	.89	-.08	.21
33	.45	.89	-.02	-.07	.45	.89	-.02	-.07
34	.45	.89	.02	-.07	.45	.89	.02	-.07
35	.45	.89	.02	-.07	.45	.89	.02	-.07
36	.45	.89	.02	-.07	.45	.89	.02	-.07
37	.45	.89	.02	-.07	.45	.89	.02	-.07
38	.45	.89	.02	-.07	.45	.89	.02	-.07
39	.45	.89	.02	-.07	.45	.89	.02	-.07
40	.45	.89	.02	-.07	.45	.89	.02	-.07

APPENDIX G (Continued)

Group WN

Trial	Experiment I 1st Rel	2nd Rel	1st Irr	2nd Irr	Experiment II 1st Rel	2nd Rel	1st Irr	2nd Irr
1	.34	.46	.40	-.01	.43	.30	.26	.00
2	.38	.43	.00	-.23	.11	.44	.21	.42
3	.20	.22	-.32	.45	.79	.35	-.06	.13
4	.50	.77	.08	-.16	.39	.64	.17	.17
5	.36	.85	.18	-.09	.45	.74	.02	.02
6	.45	.64	-.30	-.06	.50	.44	-.21	.19
7	.40	.71	.22	.01	.55	.67	-.04	-.28
8	.45	.83	-.01	-.06	.45	.89	.00	-.04
9	.50	.85	.02	-.02	.17	.81	.22	.04
10	.45	.89	-.03	-.02	.47	.86	.03	-.03
11	.50	.73	-.10	.19	.14	.80	.17	.08
12	.50	.79	.04	-.01	.39	.88	.00	-.10
13	.38	.78	-.09	.02	.25	.87	-.08	-.08
14	.40	.92	.00	-.09	.43	.86	-.10	.00
15	.40	.88	.08	.04	.50	.86	.00	-.07
16	.45	.89	.00	-.04	.39	.90	.06	-.06
17	.39	.82	-.02	.16	.38	.89	-.17	.20
18	.39	.76	-.23	.45	.40	.91	-.07	.22
19	.45	.89	-.08	.21	.49	.79	-.20	.08
20	.46	.85	-.02	.21	.37	.88	-.09	.08
21	.47	.88	-.17	.23	.42	.72	-.09	.19
22	.44	.89	-.12	.24	.39	.90	-.11	.19
23	.48	.86	-.01	.26	.52	.82	-.12	.10
24	.45	.89	-.08	.21	.35	.61	-.18	.28
25	.45	.89	-.08	.21	.19	.67	.20	.31
26	.45	.89	-.08	.21	.45	.89	-.08	.21
27	.45	.89	-.08	.21	.45	.89	-.08	.21
28	.45	.89	-.08	.21	.50	.82	-.19	.19
29	.45	.89	-.08	.21	.39	.88	-.05	.10
30	.45	.89	-.08	.21	.50	.78	-.04	.26
31	.45	.89	-.08	.21	.44	.89	-.12	.24
32	.45	.89	-.08	.21	.47	.88	-.05	.24
33	.45	.89	.02	-.07	.45	.89	.02	-.07
34	.45	.89	.02	-.07	.45	.89	.02	-.07
35	.45	.89	.02	-.07	.45	.89	.02	-.07
36	.45	.89	.02	-.07	.45	.89	.02	-.07
37	.45	.89	.02	-.07	.45	.89	.02	-.07
38	.45	.89	.02	-.07	.45	.89	.02	-.07
39	.45	.89	.02	-.07	.45	.89	.02	-.07
40	.45	.89	.02	-.07	.45	.89	.02	-.07

APPENDIX G (Continued)

Group XS

Trial	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
1	.21	.52	-.02	.04	.60	.40	-.40	-.05
2	-.10	.04	-.33	.41	.42	.44	-.30	.24
3	.42	-.05	.05	-.49	.39	.47	.43	-.07
4	.03	.03	-.20	.41	.46	.24	.10	.15
5	.45	.49	.08	.28	-.11	.70	.29	-.02
6	.12	.46	.20	-.20	.15	.59	-.08	-.19
7	.06	.87	.18	-.08	.37	.60	.32	.06
8	.34	.22	.22	-.36	.23	.43	-.33	.50
9	.10	.65	-.01	.01	.46	.35	.30	.28
10	.22	.49	.01	-.11	.43	.58	-.25	-.08
11	.45	.61	.12	-.09	.39	.68	.19	.08
12	.35	.53	-.12	-.35	.57	.44	-.23	-.02
13	.40	.30	-.16	.34	.39	.55	-.23	.31
14	.07	.72	.40	.09	.27	.74	.25	.18
15	.04	.46	-.20	.39	.33	.77	.06	-.06
16	.12	.53	.38	-.29	.29	.82	-.01	-.07
17	.39	.85	-.14	.22	.61	.61	.19	.32
18	.22	.22	.62	.17	.47	.88	-.12	.21
19	.30	.89	-.22	.10	.45	.89	-.08	.21
20	.38	.82	-.07	.40	.42	.90	-.09	.20
21	.61	.68	-.27	.18	.45	.89	-.08	.21
22	.45	.89	-.09	.17	.45	.89	-.08	.21
23	.47	.77	-.07	.21	.45	.89	-.08	.21
24	.61	.63	.16	.36	.45	.89	-.08	.21
25	.50	.85	-.09	.22	.45	.89	-.08	.21
26	.46	.88	-.09	.25	.45	.89	-.08	.21
27	.50	.78	-.04	.26	.45	.89	-.08	.21
28	.45	.87	-.01	.16	.45	.89	-.08	.21
29	.45	.89	-.08	.21	.45	.89	-.08	.21
30	.45	.89	-.08	.21	.45	.89	-.08	.21
31	.45	.89	-.08	.21	.45	.89	-.08	.21
32	.45	.89	-.08	.21	.45	.89	-.08	.21
33	.45	.89	.02	-.07	.45	.89	.02	-.07
34	.45	.89	.02	-.07	.45	.89	.02	-.07
35	.45	.89	.02	-.07	.45	.89	.02	-.07
36	.45	.89	.02	-.07	.45	.89	.02	-.07
37	.45	.89	.02	-.07	.45	.89	.02	-.07
38	.45	.89	.02	-.07	.45	.89	.02	-.07
39	.45	.89	.02	-.07	.45	.89	.02	-.07
40	.45	.89	.02	-.07	.45	.89	.02	-.07

APPENDIX G (Continued)

Group XO

Trial	Experiment I 1st Rel	2nd Rel	1st Irr	2nd Irr	Experiment II 1st Rel	2nd Rel	1st Irr	2nd Irr
1	-.22	.01	.24	.08	.16	-.06	.23	.60
2	-.02	.15	.58	-.09	-.28	.32	.44	.15
3	.16	.50	.03	-.08	-.14	.16	.11	.36
4	.01	.32	.11	-.28	.08	.42	.19	.15
5	.47	.44	.19	-.26	.48	.44	.14	.04
6	-.03	.62	-.11	-.17	-.01	.71	-.15	-.27
7	.55	.32	-.35	-.03	.42	.61	-.03	.31
8	.11	.32	.11	-.15	.21	.24	-.16	-.27
9	-.06	.44	.42	.25	.58	.46	-.13	.23
10	.53	.46	.15	.37	-.09	.42	.33	-.11
11	.48	.52	-.15	.27	.38	.63	.10	-.20
12	.46	.73	.00	-.07	.44	.67	-.15	.28
13	.40	.78	-.02	.05	.22	.71	.29	.17
14	.09	.54	-.13	-.11	.26	.67	.23	-.32
15	.48	.72	.00	.04	.36	.80	-.08	-.04
16	.33	.77	-.16	.18	.40	.86	.08	-.02
17	.17	.87	-.21	.07	.39	.81	-.02	.27
18	.58	.62	-.12	.22	.49	.82	-.09	.16
19	.30	.77	-.06	.29	.46	.64	-.16	.39
20	.51	.62	.09	.35	.38	.88	.02	.27
21	.13	.73	-.08	.34	.36	.73	-.02	.21
22	.32	.87	-.14	.17	.52	.82	.03	.24
23	.08	.64	.04	.50	.42	.77	.11	.06
24	.28	.67	.01	.29	.45	.89	-.08	.21
25	.45	.89	-.09	.22	.53	.78	.06	.24
26	.37	.74	.06	.08	.59	.70	-.14	.14
27	.33	.87	-.13	.32	.45	.89	-.08	.21
28	.50	.63	-.30	.14	.45	.89	-.08	.21
29	.41	.89	-.11	.21	.45	.89	-.08	.21
30	.49	.68	.12	.13	.45	.89	-.08	.21
31	.43	.82	.07	.09	.45	.89	-.08	.21
32	.31	.69	-.08	.03	.45	.89	-.15	.21
33	.18	.84	.00	.07	.45	.89	.02	-.07
34	.45	.89	.02	-.07	.35	.86	.08	-.12
35	.36	.75	.22	-.20	.45	.89	.02	-.07
36	.43	.90	-.01	-.06	.45	.89	.02	-.07
37	.15	.61	-.29	.03	.45	.89	.02	-.07
38	.45	.89	.02	-.07	.45	.89	.02	-.07
39	.50	163	-.20	-.29	.45	.89	.02	-.07
40	.31	.92	.08	-.12	.45	.89	.02	-.07

APPENDIX G (Continued)

Group XN

Trials	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
1	.20	-.01	-.17	.34	.18	.67	.42	-.03
2	-.06	.25	.25	.11	.31	.05	-.33	.07
3	-.30	.12	-.23	.42	.58	.18	.36	-.06
4	.18	.61	.18	-.13	.55	.60	-.09	.07
5	.57	.04	.00	-.55	.28	.17	-.20	.40
6	.25	.79	-.08	-.04	.50	.10	.26	-.12
7	.21	.37	.19	.07	.29	-.12	.19	.10
8	.48	.48	.19	.21	.74	.33	.04	.18
9	.14	.68	-.19	.19	.16	.52	-.06	-.36
10	.27	.82	.01	.06	-.13	.03	.39	-.34
11	.23	.78	-.27	-.09	.11	.45	.21	.04
12	.34	.42	.30	.38	.52	.67	-.18	.07
13	.40	.75	-.20	-.10	.41	.70	.15	-.04
14	.17	.71	-.27	-.17	.49	.70	.10	-.21
15	.58	.54	.14	.10	.61	.53	.13	-.23
16	.42	.44	.07	-.14	.50	.63	.03	-.03
17	.23	.86	.10	.38	.42	.81	.10	.19
18	.60	.64	-.19	.18	.46	.70	-.01	.27
19	.29	.82	-.05	.46	.45	.87	-.08	.13
20	.16	.56	-.24	-.13	.40	.68	.00	.16
21	.35	.85	-.17	.16	.22	.92	-.28	.08
22	.46	.87	-.02	.23	.45	.78	-.19	.06
23	.52	.80	-.08	.07	.14	.61	-.06	.53
24	.44	.89	-.07	.20	.25	.71	-.22	.06
25	.35	.86	.06	.28	.25	.56	.28	.29
26	.40	.84	-.03	.17	.23	.78	-.11	.15
27	.46	.86	-.01	.14	.55	.79	-.12	.06
28	.42	.90	-.05	.22	.52	.64	-.33	.28
29	.37	.86	-.04	.07	.46	.57	.08	.36
30	.45	.89	-.08	.21	.43	.62	.17	.09
31	.35	.61	.21	-.04	.34	.62	.17	-.01
32	.20	.77	-.30	-.01	.52	.67	-.17	.12
33	.26	.87	-.03	.09	-.20	.52	.12	.18
34	.45	.89	.02	-.07	.39	.87	-.01	.06
35	.45	.89	.02	-.07	.45	.88	.01	-.09
36	.45	.89	.02	-.07	.27	.60	.31	-.13
37	.45	.89	.02	-.07	.35	.64	.00	.13
38	.50	.78	.07	-.02	.42	.90	.01	-.08
39	.38	.92	.08	-.06	.64	.58	-.27	-.03
40	.60	.64	-.17	.12	.34	.87	.19	-.23

APPENDIX G (Continued)

Control Groups

Trial	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
1	-.12	.56	.38	-.09	.40	.70	.15	-.11
2	.34	-.13	.15	-.38	.30	.34	-.34	.03
3	.37	.03	-.18	.09	.30	-.16	-.30	.12
4	.22	.13	.62	.53	.73	.27	.12	.09
5	-.24	.02	.62	.00	-.35	.33	.16	.20
6	.44	.21	-.02	-.10	.32	.18	.04	.25
7	.62	-.32	-.16	-.03	.60	.18	-.07	.16
8	.53	-.22	-.04	.01	.27	-.23	.27	.02
9	.07	.02	.55	.43	.00	-.06	.28	.15
10	.10	.70	.34	.08	.46	.13	.08	.26
11	.08	.11	-.21	.25	.28	.36	.18	-.03
12	.27	.07	.53	.49	-.03	.71	-.07	-.10
13	.26	.16	-.26	.22	.15	.62	.13	-.04
14	-.03	.04	.14	.15	.47	.65	.19	.14
15	.24	-.29	.21	-.11	.72	.14	.30	-.18
16	-.19	.58	.39	-.06	.05	.47	.09	.24
17	.10	.52	-.13	.12	.57	.57	.08	-.09
18	.08	-.07	.43	.02	.12	.83	-.18	.38
19	.41	.47	-.04	.43	.39	.45	.17	.55
20	.52	.52	.01	.18	.02	.20	-.23	-.01
21	.49	.13	-.08	-.10	.44	.55	.11	.47
22	.31	.35	-.02	.53	.37	.68	.04	.27
23	.29	.70	-.03	-.12	.14	.28	.22	.54
24	.52	.43	-.08	.24	.18	.65	-.01	.22
25	.19	.46	.29	.14	.36	.57	.27	.11
26	.24	.70	.08	.36	.44	.67	.07	-.08
27	.44	.70	-.08	.29	.19	.83	-.27	.17
28	-.26	.70	-.08	.17	.44	.47	-.23	.20
29	.62	.51	.15	.35	.14	.54	.18	-.06
30	.32	.75	-.22	.37	.28	.72	-.11	.34
31	.69	.34	.04	.42	.33	.86	-.21	.23
32	.14	.75	-.17	.45	.33	.76	.14	.34
33	.58	.71	-.07	-.04	.43	.90	.03	-.08
34	.28	.82	-.21	.08	.41	.67	.06	.17
35	.44	.72	-.01	.18	.22	.72	.07	-.05
36	.29	.71	-.32	-.06	.13	.60	-.03	.13
37	.39	.75	.14	.18	.44	.70	.05	.00
38	.14	.77	-.03	.11	.34	.90	.11	-.05
39	.43	.61	.25	-.06	.57	.63	-.14	.01
40	.17	.24	.46	.24	.38	.79	.00	.04

APPENDIX H

Average Relevant and Irrelevant Cue Criterialities for
Overlapping Trials on the Criterion Task
(Figures 7, 8, and 9)

Blocks of Trials	Same Cue Groups		Opposite Cue Groups		New Cue Groups	
	Relevant	Irrelevant	Relevant	Irrelevant	Relevant	Irrelevant
1-2	34.00	19.44	11.75	35.06	28.88	20.03
2-3	35.88	22.94	18.19	27.88	29.81	22.81
3-4	36.94	20.00	18.50	23.69	43.62	18.31
4-5	39.25	17.75	28.50	26.88	48.12	15.69
5-6	42.44	16.69	38.56	19.31	46.44	17.00
6-7	43.25	14.94	43.56	17.94	44.12	14.75
7-8	48.12	18.44	45.38	15.19	49.81	11.69

APPENDIX I

Single S Criterialities--Block 1

Group WS

S#	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
1	.77	.30	.01	.20	.45	.30	-.19	.50
8	.49	.80	.09	-.04	.36	.90	.04	-.18
15	.45	.89	.00	-.04	.45	.89	.03	-.08
22	-.04	.10	-.13	.37	.50	.68	-.30	.01
29	.34	.77	.08	.04	.09	.15	.18	.67
36	.39	.74	-.17	.03	.41	.82	-.07	-.15
43	-.10	.62	-.03	.08	.42	.88	.01	-.04
50	.45	.89	.00	-.04	.62	.30	.20	-.28
57	.47	.86	-.02	.02	.31	.74	.24	.07
64	.45	.89	.00	-.04	.54	.50	-.02	.38
71	.48	.78	-.02	-.02	.30	.76	.06	.28
78	.24	.83	-.14	-.12	.45	.89	.00	-.04
85	.70	.21	.27	-.25	.51	.62	-.25	-.12
92	.25	.55	-.03	.17	.45	.89	.00	-.04
99	.21	.76	.27	.11	.40	.91	-.01	-.08
106	.60	.76	.00	.10	.29	.58	.42	-.08

APPENDIX I (Continued)

Group WO

S#	Experiment I 1st Rel	2nd Rel	1st Irr	2nd Irr	Experiment II 1st Rel	2nd Rel	1st Irr	2nd Irr
2	.37	.63	.11	-.20	.42	.60	.23	-.17
9	.45	.89	.00	-.04	.36	.90	.04	-.18
16	.34	.20	-.09	.26	.35	.79	.25	-.30
23	.27	.21	-.45	-.05	.16	.16	-.06	.57
30	.52	.54	-.06	.11	.23	.47	.28	.28
37	.23	.59	.18	.01	.12	.27	.17	-.15
44	-.22	.54	.58	.24	.05	.64	.43	.02
51	-.11	.54	.24	.34	.20	-.20	.68	.39
58	.22	.57	.22	.18	.11	.70	.27	-.01
65	.47	.84	-.09	.05	.39	.52	.17	.60
72	.35	.50	.33	.33	.20	.74	.14	.29
79	.45	.89	.00	-.04	.28	.68	.37	.18
86	.39	.90	.05	.01	.36	.53	.10	.12
93	.09	.76	.11	.35	.30	.66	-.08	.19
100	.21	.76	.27	.11	.32	.63	.28	.19
107	.01	.09	.41	.07	.27	.86	.10	.12

APPENDIX I (Continued)

Group WN

S#	Experiment I 1st Rel	2nd Rel	1st Irr	2nd Irr	Experiment II 1st Rel	2nd Rel	1st Irr	2nd Irr
3	.37	.61	-.13	-.24	.47	.81	.04	-.08
10	.23	.40	.02	-.25	.53	.83	-.01	.04
17	.60	.65	.00	-.02	.55	.70	-.05	.00
24	.44	.89	.03	-.05	.48	.85	.01	.01
31	.30	.46	-.22	.00	.45	.89	.00	-.04
38	.41	.89	.02	-.07	.23	.54	.26	.00
45	.35	.86	.05	-.10	.46	.88	-.01	-.01
52	.52	.56	.00	.05	.47	.33	.27	-.10
59	.25	.78	.20	.05	.42	.88	.06	.01
66	.37	.63	.09	.37	.41	.79	-.11	.11
73	.50	.85	.02	-.02	.45	.89	.00	-.04
80	.40	.88	.08	.04	.60	.71	.05	-.16
87	.42	.79	-.02	.10	.34	.92	.05	.05
94	.40	.84	-.13	.09	.52	.56	-.24	-.04
101	.45	.89	.03	-.08	-.37	.27	.15	-.07
108	.56	.63	.10	.01	.39	.39	.12	.47

APPENDIX I (Continued)

Group XS

	Experiment I				Experiment II			
S#	1st Rel	2nd Rel	1st Irr	2nd Irr	1st Rel	2nd Rel	1st Irr	2nd Irr
4	.09	-.06	.06	.25	.47	.72	-.12	-.02
11	.16	.34	.50	.07	.33	.90	.06	-.18
18	.11	.06	-.13	.20	.34	.19	-.22	.12
25	-.49	-.10	.30	.21	.49	.85	-.06	.04
32	.26	.54	.01	-.01	.38	.38	-.05	-.03
39	.06	.18	-.15	.00	.48	.78	-.12	-.12
46	.23	.72	.08	-.14	.32	.89	-.02	-.14
53	.53	.82	.08	-.02	.26	.72	-.01	.41
60	.54	.69	.16	-.10	.27	-.02	-.05	.36
67	-.07	.07	-.07	-.73	.49	.59	.04	.41
74	.42	.79	-.10	.05	.45	.89	.03	-.08
81	.08	.00	.15	-.02	.29	.75	.10	-.03
88	.32	.62	-.15	.18	.12	.21	.23	.05
95	.27	.86	-.09	.17	.13	.20	.20	.02
102	.53	.39	-.13	.31	.51	.19	.21	.27
109	.37	.84	-.08	-.12	48	.86	-.06	-.10

APPENDIX I (Continued)

Group XO

S#	Experiment I 1st Rel	2nd Rel	1st Irr	2nd Irr	Experiment II 1st Rel	2nd Rel	1st Irr	2nd Irr
5	-.30	.10	-.36	-.23	.34	.76	.03	-.17
12	.16	.22	.40	-.43	.36	.79	.00	.07
19	.44	.56	.08	-.13	.47	.71	-.12	-.15
26	.44	.89	.03	-.05	-.04	.26	-.44	.00
33	.48	.85	-.02	.02	.19	.60	.14	.15
40	.15	.47	.35	-.15	-.17	.12	.48	-.19
47	-.07	.58	.12	.12	.07	-.31	-.29	.24
54	.23	.18	.28	.14	.38	.47	.13	.46
61	.46	.34	.00	.41	.41	.89	.07	-.02
68	.38	.46	.06	.06	.11	-.07	.29	.21
75	.45	.89	.00	-.04	.39	.81	-.05	.17
82	.21	.62	-.11	.05	.45	.89	.03	-.08
89	.49	.81	-.12	-.01	.36	.92	.02	.02
96	.32	.92	.03	.05	.24	.58	.35	.19
103	-.13	.15	.36	-.17	.12	.66	.10	.31
110	.00	.32	.06	.19	-.30	.12	.66	-.16

APPENDIX I (Continued)

Group XN

S#	Experiment I 1st Rel	2nd Rel	1st Irr	2nd Irr	Experiment II 1st Rel	2nd Rel	1st Irr	2nd Irr
6	.53	.76	-.01	-.01	.27	-.11	.14	-.30
13	.22	-.08	-.35	-.03	.45	.70	.12	-.16
20	.46	.87	-.08	.03	.52	.69	-.02	.07
27	.45	.89	-.03	-.03	.45	.87	-.07	-.02
34	.11	.26	.09	.16	.41	.84	-.06	-.19
41	.26	.70	.21	.01	.30	.51	.04	-.02
48	.59	-.01	.01	.53	.45	.89	.00	-.04
55	.12	.56	.08	.33	.40	.84	.12	.09
62	.37	.84	-.10	.12	.39	.77	.07	.05
69	-.22	.11	-.16	-.10	.19	.37	.48	.06
76	.48	.72	-.03	.08	.01	-.34	-.42	.29
83	.45	.89	.00	-.04	.34	.85	-.07	.15
90	.35	.68	-.18	-.32	.60	.34	.30	-.38
97	.18	.12	.43	.01	.40	-.16	-.25	.08
104	-.10	.12	.05	-.01	.30	-.39	.50	.06
111	.20	.77	.22	.18	.32	-.11	.30	-.13

APPENDIX I (Continued)

Control Group

S#	Experiment I 1st Rel	Experiment I 2nd Rel	Experiment I 1st Irr	Experiment I 2nd Irr	Experiment II 1st Rel	Experiment II 2nd Rel	Experiment II 1st Irr	Experiment II 2nd Irr
7	-.68	-.03	.14	-.08	-.01	.74	.15	.21
14	-.34	.05	.03	-.03	.08	.42	.14	.20
21	.33	.43	.31	-.28	.25	.27	-.11	.25
28	.71	.38	.16	.24	.51	.23	-.13	.38
35	.07	.02	.46	.10	.03	-.01	.22	-.38
42	.20	.22	.60	.02	.45	.89	.00	-.04
49	.48	.30	.50	.21	-.04	.67	.18	-.41
56	.30	-.16	.43	.41	-.08	-.08	-.32	.20
65	.10	.33	.06	-.10	.47	.37	-.12	.44
70	.35	-.38	-.43	.08	.70	.48	.30	-.04
77	.30	.78	.03	.08	.12	.29	.07	-.05
84	.04	-.26	.53	.24	.46	-.35	.09	.19
91	.15	.52	-.17	.21	.66	.38	.14	.32
98	.55	.25	.16	.39	.56	.57	.25	-.09
105	.30	-.32	.40	.05	.28	-.10	.44	-.25
112	.00	-.10	.08	.02	.44	.12	.01	.12

VITA

Dale Edward Mattson was born on April 5, 1934 in Newberry, Michigan. He received his elementary and secondary education in Newberry, Michigan. He graduated from North Park Junior College in 1957 and received his Bachelor of Arts from Colorado College in 1959.

In the fall of 1960, he entered the University of Illinois as a graduate student in the College of Education; he held an N.D.E.A. fellowship in a special program in the Psychology of Classroom Learning.

He received a Master of Arts degree in June, 1961, and the Doctor of Philosophy degree in October, 1963.

I 46 p

TRAINING RESEARCH LABORATORY

Department of Psychology — Bureau of Educational Research
University of Illinois — 8 Lincoln Hall — Urbana, Illinois

LEARNING AND TRANSFER EFFECTS OF PROMPTING AND CONFIRMATION PROCEDURES

Devore Eugene Killip

Technical Report No. 2

Psychological and Educational Factors in
Transfer of Training
Phase I

October, 1963

U. S. Office of Education
Contract 2-20-003

Lawrence M. Stolurow
Principal Investigator

THE EFFECT OF DISCOVERY ON THE LEARNING OF MANUAL SKILL

BY

DEVORE EUGENE KILLIP

THESIS

LEARNING AND TRANSFER EFFECTS OF PROMPTING AND CONFIRMATION PROCEDURES[1]

Devore Eugene Killip

Technical Report No. 2

Psychological and Educational Factors in
Transfer of Training
Phase I

October, 1963

Lawrence M. Stolurow
Principal Investigator

Project Sponsor
Educational Media Branch
U. S. Office of Education
Title VII
Contract 2-20-003

[1] This study was accomplished as part of the author's Master's Thesis, The Effect of Discovery on the Learning of Manual Skill.

Table of Contents

(Continued on next page)

Table of Contents

(Continued)

Introduction

A dentist must possess a variety of highly-refined special skills if he is to treat his patients properly. Massler (1961) pointed out that manual skill is of such importance that the dental student devotes more academic time to that specialized learning than he does to the acquisition of verbal knowledge. The primary instrument involved in dental skill is the cord-driven straight handpiece. It would seem important to investigate its use in a learning situation.

In doing this, it would seem useful to conceptualize the problems of the student in terms of the abilities required in working with this instrument. Fleishman (1953) derived a classification of psycho-motor abilities through factor analysis of performance tests. His analysis provides a potentially useful basis for conceptualizing the dental students task in learning to use the cord-driven straight handpiece. The performance requirements, in these terms, would seem to involve manual and finger dexterity, combined with steadiness, coordination, and psychomotor precision. Exactly how these abilities contribute and in what proportion is not established. However, in addition, it would seem that the relevant dental skills also would include the perception and discrimination of specific visual, auditory, and tactile stimuli which guide the motor activities (Klausmeier, 1961).

This study is concerned with the learning of manual skill involving these dexterities and sensory skills. Learning as defined in this study would include three processes, each of which is measured by a different set of dependent variables: (a) the acquisition of patterns of responses to complex stimuli, (b) generalization transfer--ability

to perform the task under modified stimulus conditions, and (c) inductive transfer--ability to understand the principles and relationships involved.

The basic principle in learning is association (McGeoch and Irion, 1960).

The process of acquiring a complex manual skill would involve the formation of three distinguishable sets of associations: (a) stimulus learning, (b) response learning, and (c) associative learning (Stolurow, 1961). Each set of associates is really a skill within itself, and the sets are integrated to produce a continuous but complex skill that accomplishes a specified goal. For example, the student must first learn to discriminate the cue that is critical for a particular response. The correct response must be available for association with the appropriate cue. Typically the required response is not a natural unit, but rather a synthetic response produced by the association of separately appearing unitary response. These responses must develop into a stable sequence or time series to perform a task-relevant operation in a smooth and integrated fashion. Individual responses are chained together through the cues they produce. The component parts of a complex motor pattern are linked so that the internalized cues of the task-relevant operation will be associated with each other. Once the external cues have been discriminated and the complex response sequences have become chained, the third process--associating the cues with the newly formed responses--must be accomplished.

One primary condition for associative learning is temporal contiguity. Contiguity is the closeness in time of the cue stimulus to the desired response. If the response follows closely after perception of the stimulus, that response will become transferred to that stimulus.

It is an all important condition to be maintained in motor learning, where a great number of cues must be associated with a variety of responses before a satisfactory level of performance is attained.

The more precisely the student can identify the stimuli that serve as cues and the critical stimuli can be controlled through temporal scheduling, the more efficient will be the learning situation. Verbal stimuli serve as eliciting stimuli in the typical classroom situation. If properly introduced, they can be contiguous with task cues and the responses they elicit can effectively transfer to the task cue. Lectures and demonstrations are conventional methods of directing the learning of skills in general and in particular those taught in a dental class. In this manner the instructor provides verbal stimuli that elicit the desired responses. The effective use of these stimuli depends upon their contiguity with task cues. In short, they are to be used in a way that will decrease the interval between cue and response. Secondly, they should be chosen to insure the correspondence of the elicited response with the response demands of the task. The better the match, the faster the learning.

Problem. The teaching of a manual skill involves the manipulation of the stimulus situation in such a way as to promote the development of contiguity of the essential stimulus elements in the task with the required responses. The question raised by this study is whether it is possible to structure the learning situation in such a way to more effectively promote the requisite contiguity. Two methods are possible. One method is conventional verbal instruction provided before learning trials are attempted. The other is the discovery method. With it non-guided trials are provided the student or, in other words, he is permitted

an attempt to solve the problem without explicit instruction. Degrees of discovery can be specified, but essentially they are alike in that the student is allowed freedom to establish relationships using his own resources. In both methods verbal instruction may be used. The language employed can be seen as having two roles: (a) prompting, in which the instruction precedes execution of the overt response by the learner, and (b) confirmation, in which the learner discovers the response on his own, and the language is used to verify the correctness of that response. Both uses of language would promote contiguity, but in the former situation, it is contiguity of cue and response, whereas in the latter case, it is contiguity of response and reinforcement. The question to ask is if the difference in contiguity resulting from the use of these methods in the learning of perceptual motor associations is related to different rates of acquisition, different degrees of generalization or different degrees of inductive generalization.

Hypothesis.

The hypotheses tested were: (a) prompting will lead to more rapid learning to conventionally perform a task as taught, and (b) confirmation will lead to more efficient learning in generalization and inductive transfer.

Method

Design

This was a simple two-group design. A class of students was divided into two comparable groups on the basis of their first and second quarter technic grades. The independent variable was temporal relationship between the verbal instruction and the overt practice.

Procedure.

One group was given an illustrated lecture on the use of the dental handpiece and directions on the expert method to accomplish the cutting of a prescribed amount of dentin-like material from a simulated tooth. They were then allowed one hour to practice using the handpiece on four trials of the learning task. They then were required to perform a transfer task.

The second group of students was allowed to practice for 30 minutes on two trials of the learning task given only a statement of the problem. They then received an identical lecture as the first group, after which they practiced for another thirty minutes on two additional trials of the learning task. They were then required to perform the same transfer task as the first group.

The dependent variables are motor performances during the training and transfer tasks. The controlled variables include: (1) content and method of instruction, (2) length of instruction, (3) visual illustrations, (4) length of time for motor learning, (5) amount of feedback from motor performance, (6) amount of prior technic experience, (7) previous performance in technic tests, (8) precision of measuring equipment, and (9) testing and testing conditions.

Subjects.

The Ss for this study were the freshman class of 84 students at the University of Illinois College of Dentistry. The Ss all had the same introductory use of the dental handpiece in technic courses, but no experience relating to its use as investigated in this study.

Learning Task.

The Ss of this study were required to operate a dental straight handpiece using a Number 560 straight fissure bur to remove a specific amount of dentin-like material from a simulated tooth-like area. The task was performed on four separate blocks of identical construction with a time limit of fifteen minutes imposed on each trial.

Learning Materials.

The materials used in the learning task were designed specifically to meet the need of this study. In order to simulate tooth structure, the materials were required to respond to manipulation similar to vital structure, be of corresponding size, color, and texture, yet be identically produced in quantity. The material to be cut by the bur was obtained from the Columbia Dentoform Corporation of New York as "Ivorine." This material, the closest synthetic to tooth dentine, was prepared in the shape of cylinders 1/4 inch in diameter and 1/4 inch long.

The material used to surround the prepared plugs was pure lead, selected for its soft, tough character, yielding an easily distinguished differentiation in cutting quality. To prevent the student from using any short cuts to remove the dentin-like material, the plug was cemented to the lead receptacle with epoxy cement. The simulated tooth was then

set into a white plaster diamond-shaped block, the only purpose of which was to facilitate handling during use.

The completed learning material (Figure 1) appears simple and unsophisticated, but once attempted presents a difficulty level sufficient to challenge even the expert. The design permits both visual and tactile feedback during operation, and knowledge of results upon completion of trial. Errors committed beyond the prescribed limits are recordable and measurable. The manner in which the task was performed can be observed from the finished product.

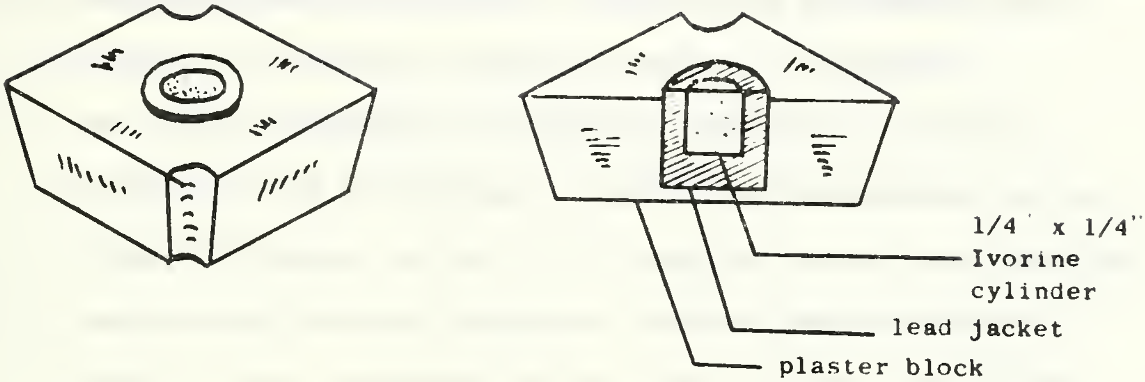


Figure 1

Learning Task Materials

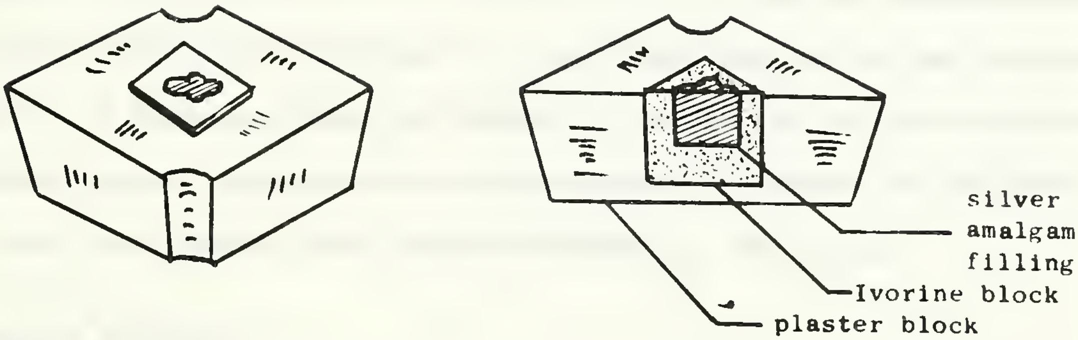


Figure 2

Transfer Task Materials

Drawn 1-1/8 actual size

Instructions

The following instructions were read to Ss prior to learning: "You have been given four white plaster blocks containing lead cylinders into which is set an ivorine or dentin-like plug 1/4 inch in diameter and 1/4 inch deep. You are to remove all the tooth-like material, taking as little of the surrounding lead material as possible. You are to use only the handpiece and Number 560 straight fissure bur. You will be allowed 15 minutes to work on each block. You will not be graded on any of these four blocks, but you will be required to carry out a similar task on which you will be graded. You have been given a plug gauge 1/4 inch in diameter and 1/4 inch in length. You may use it after you have finished each trial block to determine your accuracy of performance."

Transfer Task.

The Ss of this study were required to operate a dental straight handpiece using a Number 560 straight fissure bur to first remove a silver amalgam filling from a block of dentin-like material, and then prepare a cavity 1/4 inch in diameter and 1/4 inch deep. One test of the transfer task was a 15-minute timed performance. A second test was the identical task not timed. Each group was counterbalanced to offset any influence of the first transfer task on the performance of the second.

Transfer Materials.

In order to simulate tooth structure, and to present a challenge found under clinical conditions, the dentin-like synthetic material was prepared in 1/2 inch cubes. Each cube was cut with a cavity representing

a typical occlusal preparation in a lower molar tooth. The cavities were cut with a Number 560 straight fissure bur, identically produced, and held to a tolerance of .001 inch. The cavity so prepared measured .245 inch in one diameter, .187 inch in a second width perpendicular to the first, and .249 inch in depth. The cavity was then filled with silver amalgam as used in regular tooth restoration. Each ivorine cube was set into the same type plaster block as described under training task materials.

The design of the materials (Figure 2) presents a challenge not before encountered, yet relating to the same principles involved in the learning task. The shape of the silver filling is such that the careful student could remove it and prepare a cavity within the prescribed dimensions.

Instructions

The following instructions were read to the Ss, as well as enclosed on a printed sheet with their transfer task materials: "You have been given two white plaster blocks containing an ivorine or dentin-like block into which is set a silver amalgam filling. The filling measures a little under 1/4 inch at the widest width and in depth. You are to remove all the silver filling, for if any amalgam is present the preparation will be considered as a zero. Then you are to prepare a round cavity with a flat bottom 1/4 inch in diameter and 1/4 inch deep. You may use the plug gauge to judge the correct size as you work. You are to use the same equipment as that used in the learning task. One test will be timed to 15 minutes; the other test you may work as long as you wish. You will be graded equally on both."

Learning and Testing Conditions.

The two groups of Ss were scheduled for the same time, and kept separate during the entire period. Due to available facilities, the conventional group (considered as the control for the study) was increased to N = 45 while the experimental group was reduced to N = 39. The resulting technic grade point averages are reported in Table 1.

TABLE 1

GRADE POINT MEANS AND STANDARD DEVIATIONS[a]

Group C	Mean	SD
Control	3.01	.502
Experimental	2.98	.455

[a]An A grade is counted as 5, B as 4, C as 3, etc.

The grades represent those obtained during the three quarters of the first year technic courses.

The 15 minute lecture was tape recorded for identical instruction to both groups. Fifteen slides were projected during the lecture to illustrate the important principles. The slides were close views in color of the actual demonstration presented in a manner to yield maximum visual communication (Roshal, 1961). When possible a color slide was followed by a black and white, cross-sectional drawing of the principle involved. Since the instruction was given at a different time sequence to the two groups, an innocuous film relating to dentistry but not to the task was shown to the opposite group to balance the time interval.

To negate the effect of authority influencing performance, the assigned instructors were not present during the experiment. Instead each group session was supervised by unfamiliar proctors of similar age and experience. Motivation was emphasized by declaring the test results to be a component of the student's grade in his course of oral anatomy. The groups seemed equally involved in the experiment, as it was their first opportunity to use the dental handpiece.

To obtain a verbal measure of the effect of instruction, both groups of Ss were given a five item, free-response questionnaire and asked to state how they would describe the five most important principles of the task, listing the most important first and the least important fifth, ranked in that order. The test was answered by both groups after they had been given the lecture and two opportunities to practice the learning task. The difference between groups was the relationship of the time of instruction to their practice trials.

Evaluation of Performance.

Test evaluation was made by both subjective and objective methods. Subjective evaluation was obtained from a panel of three experts from the Department of Operative Dentistry. They separately graded the preparations for neatness, smoothness, and accuracy of cutting on an eight item Evaluation Check List for Judge's Ratings (Appendix A). The manner of grading was a three point scale, 70-80-90, in current use by the Department. Each item on the check list received one of the three number grades, and these were averaged to give a differentiating score on both the timed and untimed test. The three judge ratings were then averaged for each test to give a judge score. The reliability of evaluation by the three judges was

established using the Kendall coefficient of concordance and correcting for tied ranks. The interjudge reliability can be expressed as $\underline{W}$ = .88 on a random sample of 40 sets of judge scores.

The four learning trials were rated only subjectively using a seven item check list (absence of silver amalgam as one item). The grading was accomplished on a four-point scale, 60-70-80-90, due to the wide variation between successive learning trials. The judges' ratings were averaged to obtain a learning score for each of the four trials.

Objective evaluation, considered difficult in freehand operations such as dentistry, was thought necessary for the purpose of this experiment. To accomplish objective measurement, the work samples had to be held to .001 inch tolerance and then measured within the same limits. Since the dimensions were exactly described to the <u>Ss</u>, the first evaluation would be a maximum error measurement. This was obtained as the widest diameter measured with a Starrett small hole gauge, and the greatest depth measured with a modified Starrett depth micrometer. Accuracy was held within .001 inch. To establish error the obtained measurements were subtracted by the prescribed dimensions of .250 inch in diameter or .250 inch in depth. These two values were multiplied to give the maximum error score. In case the student failed to obtain the prescribed dimensions, that measurement was subtracted from .250 inch to give the error.

A second objective evaluation was made on the amount of material removed beyond the prescribed dimensions. This value was obtained by filling the prepared cavity with mercury, leveling the meniscus with a glass slide, and weighing the contents on a beam balance, accuracy .001 gram. The weight of mercury was subtracted by the weight of that in the true preparation, established at 3.210 grams. This gave an error score for the total amount of overcutting. No distinction was made between an error in overcutting or undercutting.

Results

In summarizing results the data obtained during the learning trials will be presented first and those obtained from the transfer task will follow.

Learning.

Mean effects between groups. Analysis of evaluation of the four learning trials (Table 2) indicates a difference in performance between the two groups. The difference was significant at the 1% level of confidence during the first three trials, and significant at the 5% level during the fourth trial. These data support the first hypothesis--prompting produced faster learning to perform a task as taught.

TABLE 2

MEAN COMPARISONS LEARNING TASK

Judges' Evaluation	Control group N = 45		Experimental group N = 39	
	Mean	SD	Mean	SD
Trial 1	75.5 **	7.30	69.0 **	5.34
Trial 2	79.5 **	5.43	73.5 **	6.11
Trial 3	80.9 **	5.23	75.2 **	6.02
Trial 4	80.8 *	6.07	77.6 *	6.42
Average of four learning trials	79.2 **	4.71	73.8 **	4.93

* Significant difference at 5% level of confidence.

** Significant difference at 1% level of confidence.

Mean effects within groups. A comparison of each performance trial within groups indicates that the experimental group learned in a progressive series of trials, each differing from the former with a significance at the 10% confidence level. This comparison applies only to the difference

between the first and second trials of the control group, for the second trial of that group was not significantly different from its third or fourth. The first and second trial of the conventional (control) group showed no significant difference from the third and fourth trial of the experimental group. Learning task ratings were made without a time score, as the student was encouraged to use all the time allowed on the learning task.

TABLE 3

MEAN COMPARISONS TRANSFER TASK: TIMED AND UNTIMED TESTS

Method of evaluation	Control group (N = 45)		Experimental group (N = 39)	
	Mean	SD	Mean	SD
Timed Test				
Judge 1	77.2	5.12	77.6	4.32
Judge 2	77.1	4.87	77.3	4.68
Judge 3	78.0	5.52	77.8	4.88
Judge Score	77.4	4.83	77.6	4.36
Maximum Error[a]	937.66	851.16	888.53	878.11
Volume Error[a]	.514	.324	.437	.258
Time Score	14.02*	2.64	12.79*	2.75
Untimed Test				
Judge 1	79.1	4.55	77.3	4.03
Judge 2	78.3	5.21	78.6	4.96
Judge 3	79.2	5.35	77.8	4.48
Judge Score	78.8	4.79	77.9	4.16
Maximum Error[a]	680.53	523.57	650.79	623.25
Volume Error[a]	.431	.254	.434	.286
Time Score	28.22**	5.14	18.84**	6.77

*Significant difference at 5% level of confidence.

**Significant difference at 1% level of confidence.

[a]Error scores inversely related to desired performance.

TABLE 4

MEAN COMPARISONS TRANSFER TASK: AVERAGE OF BOTH TESTS

Method of evaluation	Control group		Experimental group	
	Mean	SD	Mean	SD
Average of both tests				
Judge 1	78.16	4.27	77.5	3.61
Judge 2	77.69	4.55	77.9	4.27
Judge 3	78.61	4.92	77.8	4.13
Judge Score	78.15	5.11	77.8	4.56
Maximum Error[a]	809.09	687.36	769.66	750 68
Volume Error[a]	.472	.289	.435	.272
Time Score	21.12**	3.16	15.82**	4.35

*Significant difference at 5% level of confidence.

**Significant difference at 1% level of confidence.

[a]Error scores inversely related to desired performance.

Generalization Transfer.

The transfer task data are shown as performance during the timed and non-timed tests (Table 3), with each test evaluated by four different measurements. Mean scores for both tests are also presented. The comparison of the two groups using combined scores is presented in Table 4.

Mean effects: *t* test. The experimental group performed with no significant difference from the conventional group measured by judges' ratings, maximum error, and volume error scores.

A marked difference appeared in the time score for both tests, the *t* test indicates a difference between groups significant at the 1% level of confidence, and in practice resulted in an average savings in time. These data support the second hypothesis--confirmation produced more efficient learning in generalization transfer.

Matched pair frequencies: sign test. In comparing mean test scores, it is possible that small but consistent differences between matched pairs of subjects would not be detected. The number of cases in which the experimental subject scored higher than his matched control permitted investigation. Using the sign test with Chi-square approximation, $\underline{z}$ values were obtained for each of the 12 measures summarized in Table 5. These results indicate that significantly more of the experimental Ss obtained lower time scores on their performance than did their matched controls. These data for both the timed and untimed test also support the second hypothesis.

Analysis by ability levels. Test performance within three ability levels, based on first year technic grades, are shown in Tables 6 and 7. With the timed tests two significant differences were obtained for the lowest ability group. One was a lower volume error score for the experimental group. The other was a lower mean time score (Table 6).

All three levels of the experimental group demonstrated significantly lower time scores on untimed test (Table 7). From the results summarized in Table 7, the experimental treatment was not found to influence differentially any of the three ability levels used.

TABLE 5

MATCHED PAIR COMPARISON OF PERFORMANCE ON TRANSFER TASK USING SIGN TEST

Evaluation	Frequency of positive performance	N	z value
Timed Test			
Judges' Score	21	39	.47
Maximum Error	24	39	1.42
Volume Error	25	39	1.74
Time Score	30	39	3.32**
Untimed Test			
Judges' Score	21	39	.47
Maximum Error	24	39	1.42
Volume Error	21	39	.47
Time Score	33	39	4.27**
Average of Both Tests			
Judges' Score	21	39	.47
Maximum Error	24	39	1.42
Volume Error	24	39	1.42
Time Score	32	39	3.95**

**Significant difference at 1% level of confidence.

Inductive Transfer.

Principles of task. When asked to describe the five most important principles of the learning task, 50% of the control group responded correctly, whereas 57% of the experimental group responded correctly. Applying the Chi-square test to these frequencies of response, we cannot reject the null hypothesis. The two groups perceived the important principles of the task.

Determination of responses. When describing the task, the experimental group brought in free response elements not covered in the lecture in 15% of the replies. The control group on the other hand,

brought in only 2% such responses. Applying Chi-square to these frequencies, the groups were found to differ significantly at the 1% level. In other words, the experimental group verbalized other than what they were told. This implies that the view of the task held by the control Ss was more directly determined by what they were taught in the lecture. Such difference was evidenced by the content of the responses, where the control group was oriented toward following directions, grading of the project, and emphasizing errors. The experimental group seemed directed toward reporting about accuracy of their work, understanding properties of materials, and critical evaluation by the student.

TABLE 6

COMPARISONS OF TRANSFER TASK MEANS (TIMED TEST) WITHIN THREE TECHNIC GRADE LEVELS

Method of Evaluation	Control group (N = 15)		Experimental group (N = 13)	
	Mean	SD	Mean	SD
Judge Score				
Upper Third	79.4	5.52	79.9	5.35
Middle Third	78.0	5.11	77.2	3.98
Lower Third	74.8	3.81	75.7	3.67
Maximum Error				
Upper Third	678.13	527.04	427.15	198.63
Middle Third	844.60	806.48	1102.30	1227.20
Lower Third	1290.26	1064.93	1136.15	741.45
Volume Error				
Upper Third	.357	.267	.393	.267
Middle Third	.551	.359	.528	.331
Lower Third	.635**	.295	.392**	.129
Time Score				
Upper Third	13.8	2.19	13.2	3.11
Middle Third	13.1	2.26	12.7	2.59
Lower Third	15.0*	3.17	12.4*	2.69

*Significant difference at 5% level of confidence.

**Significant difference at 1% level of confidence.

TABLE 7

COMPARISONS OF TRANSFER TASK MEANS (UNTIMED TEST)
WITHIN THREE TECHNIC GRADE LEVELS

Method of Evaluation	Control Group		Experimental group	
	Mean	SD	Mean	SD
Judge Score				
Upper Third	81.3	4.59	80.1	5.11
Middle Third	78.8	5.07	78.0	3.64
Lower Third	76.4	4.49	75.7	3.62
Maximum Error				
Upper Third	395.33	281.65	460.38	363.30
Middle Third	865.60	616.67	617.07	449.62
Lower Third	780.66	517.06	874.92	894.99
Volume Error				
Upper Third	.307	.215	.406	.269
Middle Third	.512	.308	.413	.251
Lower Third	.474	.194	.485	.346
Time Score				
Upper Third	29.7**	5.17	20.38**	5.31
Middle Third	26.8**	5.68	16.30**	4.87
Lower Third	28.0**	4.42	19.84**	9.09

* Significant difference at 5% level of confidence.

** Significant difference at 1% level of confidence.

Analysis of timed and untimed tests. The transfer task was designed to obtain data from student performance in an altered stimulus situation both when time was restricted, and when it was not. Derived from empirical methods used for many years in teaching dental students manual skill, it was thought that restricting the student's time would diminish his performance on such a complex work sample task. It was, therefore, felt that students would perform better when given an unrestricted amount of time. Using the performance of all students and comparing judges' ratings between timed and untimed tests (Table 8), there was no significant difference (t = 1.25). Comparing the amount of overcutting likewise disclosed no difference (t = 1.07). There was a significant difference at the 2% level of confidence on the maximum error score (t = 2.19) and on the time score.

Correlations. The reliability of the judges' scores was .881. Correlations between judges' ratings and objective measurements varied from .528 to .675 on the timed test, .630 to .653 on the non-timed test (Table 9). Correlation of test evaluations with technic grade point average (Table 10) found judges' ratings to correlate best on the non-timed test (r = .391). The experimental group, using judges' ratings on the non-timed test gave the highest prediction of r = .404. Both maximum error and volume error scores gave less prediction than judges' ratings. Correlation of learning trials to grade point average gave little if any predictive value.

TABLE 8

ANALYSIS TIMED AND NON-TIMED TRANSFER TASK
(N = 84)

Score	Timed Test		Non-timed Test	
	Mean	SD	Mean	SD
Judges' Ratings	77.70	4.89	78.44	4.81
Maximum Error[a]	915.85[b]	851.80	666.72[b]	568.67
Volume Error[a]	.478	.296	.432	.268
Time	13.4[b]	2.75	23.8[b]	7.55

[a]Error score inversely related to desired performance.

[b]Significant difference at 2% level of confidence.

TABLE 9

CORRELATION OF EVALUATION METHODS
(N = 84)

Timed Test Transfer Task	
Judges' rating with maximum error[a]	-.675
Judges' rating with volume error[a]	-.528
Maximum error with volume error	.637
Non-timed Test Transfer Task	
Judges' rating with maximum error[a]	-.630
Judges' rating with volume error[a]	-.653
Maximum error with volume error	.759

[a]Error scores inversely related to judges' scores.

TABLE 10

CORRELATIONS OF TEST AND LEARNING TASK PERFORMANCE WITH TECHNIC GRADE POINT AVERAGE

Type of performance	Control group (N = 45)	Experimental group (N = 39)	Total Ss (N = 84)
Timed Test			
Judge 1	.208	.291	.239
Judge 2	.401	.269	.343
Judge 3	.232	.306	.262
Judge score	.296	.307	.300
Max. error[a]	-.206	-.265	-.231
Volume error[a]	-.249	.055	-.126
Time	-.139	.233	.034
Non-timed Test			
Judge 1	.317	.309	.314
Judge 2	.387	.425	.402
Judge 3	.374	.374	.375
Judge score	.380	.404	.391
Max. error[a]	-.290	-.332	-.303
Volume error[a]	-.219	-.229	-.223
Time	.100	.000	.059
Learning Task			
Trial 1	-.068	.279	.067
Trial 2	.341	.065	.202
Trial 3	.142	.107	.127
Trial 4	-.010	-.119	-.048

[a] Error scores inversely related to desired performance.

Discussion

This study indicates that varying the time at which instruction was given was related to differences in learning. One group was given the task with only a statement of the problem--a challenge to the student which caused him to attempt to identify critical characteristics of the task. Learning occurred under these conditions and could be described as discovery learning. Presumably, the students extracted perceptual cues from the task and these were subsequently associated with the verbal cues of instruction when it was given later on. Thus the instruction which followed the experience of working alone to discover ways of working can be described as a confirmation sequence. The information provided by the instruction confirmed correct discoveries. The verbal instruction preceded learning trials. With this sequence the instruction can be considered as prompting. In other words the verbal instruction can be thought of as providing prompts for the students which would increase the likelihood of correct overt response in the learning trials. From this point of view, the difference in performance on the training task can be attributed to the superiority of temporal contiguity produced by the prompting sequence. In other words, the prompting sequence produced more rapid learning because the correct responses followed more immediately the task cue than they did with the confirmation sequence.

The results of the transfer task disclosed that in quality of work the experimental group performed as well as the control group, but in significantly less time. These data indicate the superiority of the confirmation condition overprompting for generalization transfer. The confirmation sequence

appears to be more efficient in providing a basis for generalization since the student not only learns the correct associations which are also taught in the prompting sequence, but also he has the experience of extinguishing incorrect associations.

Temporal contiguity appears to be the critical mechanism for learning as revealed by significant improvement in time of performance. The control group established their associations through verbal cues more rapidly than the experimental group who relied on perceptual cues derived from task materials under conditions that forced them to discriminate among competing cues but did not also produce as contiguous conditions for cue-response relationships. The Ss learning by confirmation identified more precisely the stimuli that served as cues and elicitors, and responded to a wider range of stimuli. Thus, while learning more slowly they nevertheless transferred more effectively. The data suggests the value of prompting in a learning situation and confirmation in transfer where the skill is a form of problem solving requiring cue discrimination.

The implications from this study suggest the value of confirmation over prompting in the learning of manual skill in which problem solving ability is valued over rapid mastery of a specific set of skills. However, complex, rather than simple teaching methods, are required to accomplish the desired objectives of rapid learning and wide transfer. It is desirable to structure the learning situation in such a way to make the probability of the correct response as high as possible during the early stages of manual learning. Prompting should be applied during the early stages for rapid learning of the conventional pattern. Once the conventional motor pattern becomes relatively probably, the prompts should be withdrawn. Confirmation procedures should then be used so that the response can effectively become transferred to the

internalized cue, and then the cue is used alone. This permits the response to become attached to the cue stimulus in the most efficient manner to make the connection more effective for both acquisition of skill as taught and as generalized and inductive transfer.

Summary

Two groups of subjects were asked to learn the same motor task of removing a specified amount of dentin-like material using the dental handpiece. One group received instruction in how to attack the problem before starting to practice; the other group was given a chance to practice before receiving instruction. The hypothesis was that (a) prompting, in which the instruction preceded the overt response by the learner, would lead to more rapid learning to perform a task as taught, and (b) confirmation, in which the learner discovers cue-response relationships and uses the language of instruction to verify the correctness of his responses, would lead to more efficient generalization and inductive transfer.

The results of the four learning trials disclosed a significant difference between the two groups. These data supported the first hypothesis.

When tested on a transfer task, both timed and non-timed, the two groups performed with no significant difference according to judges' ratings, maximum error scores, and volume error scores. There was a significant difference at the 1% level of confidence on the time of performance. These data support the second hypothesis.

Bibliography

Fleishman, A. E. Testing for psychomotor abilities by means of apparatus tests, _Psychol. Bull._, 1953, 50, 241-262.

Klausmeier, H. J. _Learning and human abilities_. New York: Harper, 1961, 226-253.

McGeoch, J. A. and Irion, A. L. _The psychology of human learning_. New York: Longman, 1960, 43-56.

Massler, M. In M. Massler and A. Watson (Eds.), _Teaching and learning manual skills_, _Proceedings of a faculty workshop_, Chicago: University of Illinois, 1961. Pp. 2-3.

Roshal, S. M. Film-Mediated learning with varying representation of the task: Viewing angle, portrayal of demonstration, motion, and student participation. In A. A. Lumsdaine (Ed.) _Student response in programmed instruction_, _Nat. Academy of Sci. - Nat. Res. Council_, Publication 943, Washington, D. C., 1961. Pp. 155-176.

Siegel, S. _Nonparametric statistics_. New York: McGraw, 1956, 213-219.

Stolurow, L. M. Learning manual skills, _J. Dental Educ._, 1961, 25(4). Pp. 308-317.

Appendix A

Evaluation Check List for Judges' Ratings

Learning Task

1. Overall general appearance of neatness, smooth cutting, accuracy of size, size, and sharpness of surface edges.

2. Amount of error evidenced on top surface of block.

3. Amount of material remaining on side walls of cavity.

4. Amount of material remaining on floor of cavity.

5. Degree of accuracy of cavity side walls, smoothness, roundness, parallel walls, and condition of angle between floor and walls.

6. Nicks or gouges on sidewalls or floor, indicating error.

7. Gauge inserted and tested for degree of looseness of fit.

Transfer Task

1. Overall general appearance of neatness, roundness of cavity, sharp edges, alinement of walls, smoothness of floor, and size.

2. Amount of error evidenced on top surface of block.

3. Amount of amalgam filling remaining.

4. Degree and amount side walls are paralleled.

5. Degree of roundness of preparation.

6. Degree floor is level and flat.

7. Gauge inserted to determine depth of preparation.

8. Gauge inserted and tested for degree of fit.

TRAINING RESEARCH LABORATORY

Department of Psychology Bureau of Educational Research

University of Illinois 8 Lincoln Hall Urbana, Illinois

LEARNING AND TRANSFER EFFECTS OF PROMPTING AND CONFIRMATION PROCEDURES

M. David Merrill

TECHNICAL REPORT NO. 3

Psychological and Educational Factors in
Transfer of Training
Phase I

November, 1963

U. S. Office of Education
Contract 2-20-003

Lawrence M. Stolurow
Principal Investigator

TRAINING RESEARCH LABORATORY
Bureau of Educational Research
Department of Psychology
University of Illinois
Urbana, Illinois

LEARNING AND RETENTION EFFECTS OF A MODEL AND A PREVIEW IN TEACHING AN IMAGINARY SCIENCE#

M. David Merrill

TECHNICAL REPORT NO. 3

November, 1963

Psychological and Educational Factors in Transfer of Training
Phase I

Principal Investigator:
Lawrence M. Stolurow

Project Sponsor:
Educational Media Branch
U.S. Office of Education
Title VII
Contract # 2-20-003

#Submitted in partial fulfillment of the requirements for the degree of Master of Arts in Education in the Graduate College of the University of Illinois, 1964.

ACKNOWLEDGMENT

The author wishes to acknowledge the invaluable assistance received from Carl Bereiter, whose imaginary science was used as learning material; from David Ausubel, whose theoretical ideas form the basis of this study; and from Lawrence Stolurow, whose suggestions greatly assisted in preparation of the teaching program and writing of this report.

A very large thank you is also extended to Superintendent Nicholas and Director of Special Services Irvin Monash of LaDue School District for their assistance in providing subjects and to Richard F. Stauffer, Charles Cain, Thomas W. Penn, Helen Dieterich, Dean Gray and the students of Hortin Watkins High School without whose assistance and help collection of the data would have been impossible.

Table of Contents

Learning and Retention Effects of a Model and a Preview in Teaching an Imaginary Science

M. David Merrill

Piaget has suggested that as the child grows intelligence develops gradually through a series of stages each characterized by a particular way of interacting with the environment. He has further hypothesized that those events which are most conducive to maximal development at one stage are not necessarily the events most conducive at later stages. At about 11 or 12 years until about 15 or 16 a child passes from a stage of concrete operations, during which the most important aspect of his environment is that he have a wide variety of objects, gadgets, materials, and experiences with which he must cope and understand, to a stage of formal operations during which the most important kind of learning experience is verbally presented abstract arguments and conceptions.

It is with the question of how to effectively present complex verbal materials to high school students that this paper is concerned. Some hypotheses for doing this come from Ausubel (1962a) who has formulated a Subsumption theory. He proposes that the cognitive organization of the human mind "... consists of a hierarchical structure in which the most inclusive concepts occupy a position at the apex of the structure ... (and that under these concepts are subsumed) ... progressively less inclusive and more highly differentiated subconcepts and factual data." (Ausubel, 1962b, p. 651). It is also assumed by subsumption theory that not only is knowledge represented by a hierarchical organization, but that acquisition of new knowledge follows the same pattern. That is, the student first

learns general concepts and then subsumes new concepts and facts under the relevant and most appropriate concepts which exist in his cognitive structure (Ausubel, 1962a, 1962b, 1962c).

Ausubel (1962a) is careful to specify that his subsumption theory applies only to meaningful verbal reception learning and that different principles are probably involved in the acquisition of other kinds of learning, such as discovery learning, conditioning, learning of motor skills, etc. Reception learning refers to the distinction made on the basis of " ... whether the content of the learning task (i.e. what is to be learned) is presented to or independently discovered by the learner." (Ausubel, 1962a, p. 213). "By 'meaningful learning' we ... refer ... to a distinctive kind of learning process ... (which) presupposes, in turn, both that the learner employs a meaningful learning set and that the material he learns is potentially meaningful to him. Thus, regardless of how much potential meaning may inhere in a given proposition, if the learner's intention is to memorize it verbatim, ... both the learning process and the learning outcome must necessarily be rote and meaningless. And conversely, no matter how meaningful the learner's set may be, neither the process nor outcome of learning can possibly be meaningful if the learning task itself consists of purely arbitrary associations as in paired-associate or rote serial learning." (Ausubel, 1963, p. 21-22). The potential meaningfulness of learning material is determined by two criteria: "... nonarbitrary relatability to relevant concepts in cognitive structure, ... (and) its relatability to the particular cognitive structure of a particular learner ..." (Ausubel, 1963, p. 23).

Subsumption theory states that acquisition of adequate cognitive structure depends on two factors. First, the using for organizational and integrative purposes those substantive concepts and principles in a given discipline that have the widest explanatory power, inclusiveness, generality, and relatability to the subject matter content of that discipline. And second, the using of methods for presenting and ordering the sequence of subject matter that best enhance the clarity, stability, and integratedness of cognitive structure for purposes of new learning and problem solving. Ausubel feels that the principles of progressive differentiation and integrative reconciliation implemented by the use of advance organizers provide guidance to the teacher or programer of educational materials. In operational form, these principles and the use of advance organizers can be stated as follows:

Progressive differentiation: (a) Present the most general and inclusive concepts first. (b) Then, present details and specific concepts, always proceeding from general to specific.

Integrative reconciliation: (a) Point out relationships between related ideas as they come up in various divisions of the discipline. (b) Point out significant simularities and differences between ideas previously learned and the new ideas to be learned. (c) Reconcile real or apparent inconsistencies between previously learned concepts and the new concepts.

Ausubel points out that one way to implement these two principles is by the use of advance organizers. One should: (a) Present the organizer in advance of the learning material itself. (b) Formulate the organizer in terms and concepts already familiar to the learner. (c) Formulate the organizer in terms that are at a higher level of generality and inclusiveness than the terms used to present the learning material. (d) Select the

organizer on the basis of its suitability for explaining, integrating, and interrelating the material it preceeds. This means the organizer should point out ways that previously learned, related concepts are either basically similar to, or essentially different from, new ideas and information in the learning task. The organizer should also point out how the concepts to be learned are interrelated to one another. (e) For complex subject matter, use a hierarchical series of organizers (Ausubel, 1962b).

It is perhaps possible to distinguish between different types of organizers. A pedagogic practice which has been used for many years and which is related to advanced organizers is the use of models to summarize a large number of complex facts. A model in this context would be described as a shorthand way of representing complex materials by using a few general concepts which the students already understand. A model may also have the unique characteristic of using analogous concepts to summarize the facts under consideration. A model of this type, if presented prior to the learning of some material related to the model or capable of subsumption under the concepts of the model, would meet the requirements of an advance organizer.

Another type of advance organizer would be a typical summary presented prior to the learning of the material. "How to study" courses often suggest that the student read the headings of the chapter and the chapter summary before beginning to read the text material as a guide to what is important and what is discussed in the material (see Robinson, 1946). If it is profitable to read the summary first than perhaps it would be profitable to write the summary first thus assuring that it would be read in this position. Murphy (1962) demonstrated that presenting the

summary first leads to better learning and retention. Such a summary presented before the learning material itself will be called in this paper a preview. A summary presented after the learning material itself will be called a review.

The subject matter chosen for this study is a complex imaginary science called the Science of Xenograde (Zĕn ō grāde) Systems.[1] This subject matter was chosen for the following reasons: First, this science requires several hours to learn and yet it is a complete system. Second, the science is composed of several logical divisions of material (see description of learning material below) rather than only one main topic as has previously been the case in studies of this sort (Ausubel, 1960; Ausubel and Fitzgerald, 1962; Murphy, 1962). This allows one to evaluate the effect of an organizer that covers a broader range of material than a single short passage and allows the material to be organized into logical divisions as is typically done in text books and in course lecture presentations. Third, the nature of the science allows one to more easily test for comprehension rather than for mere recall of verbal statements. By presenting the student with some new data and asking him to solve problems that require applying the principles he has learned, one has a better measure of the depth of the student's understanding than if he is merely asked to recall or recognize previously learned facts. Since subsumption theory requires that the student meaningfully relate new learning to prior concepts in his cognitive structure such a problem-solving

---

[1]This science was developed by Carl Bereiter, Training Research Laboratory, University of Illinois, for use in studying group interaction problems in interdisciplinary research, ONR Contract Nonr. 1834-36, "Group and Organizational Factors Influencing Creativity."

procedure would be valuable to help access the degree of meaning the material had for the students. Three types of test items were identified which allowed us to compare problem solving with verbal recall (see description of tests below). Fourth, being an imaginary science and hence composed of laws and relationships that are not derived from real life it is unlikely that any student will already have knowledge of this subject. This allows one to more easily equate groups on previous knowledge of the subject and makes it less likely that the results will be distorted by previous knowledge. If the material is adequately taught this characteristic should also have a tendency to reduce the variance since it is unlikely that some subjects will already know more about the subject and everyone should have the same opportunity to learn the new material. Fifth, this science was developed on the basis of a simple model which uses simple concepts from everyday life. This provided a ready-made example of a model which could be used in the presentation and assured us that this model would be relevant to the science and would thus meet the requirements of a good advance organizer. Since the laws of the science are logical derivations from simple premises in the model, they are therefore nonarbitrary and potentially meaningful.

The mode of presentation chosen for this material was programed instruction. This method of presentation was chosen for several reasons. First, whenever one investigates an educational question one is faced with the problem of keeping the presentation constant from situation to situation. Even when the same teacher is used to present the material, because of different types of pupils from class to class, etc., there is likely to be wide variance in how the material is presented. When the

teaching material is in the form of a teaching machine program it can be more easily assumed that the method of presentation remains constant from student to student. Second, oral presentation becomes very difficult to examine to determine where the students didn't comprehend the material being presented. When they are required to respond to questions as they are presented the material it is possible to examine their responses and determine at which points the presentation was not clear. These points of weakness can then be rewritten for future presentations. Third, if two groups are each given a presentation by a teacher and the presentations differ, it is difficult to know if the difference resulted from the variable under consideration or whether the difference resulted from the fact that one presentation was better than the other because of some unknown factors. The use of programed instruction allows one to equate the programs much more adequately. A fourth advantage of programed instruction results from requiring the student to respond to the program and giving him information as to the correctness of his response. By monitoring his responses as he proceeds through the program one can control the degree of mastery for any given sequence of items before allowing the student to progress to the next frames. This is a particularly important advantage for the principle of progressive differentiation which assumes that previous steps are clear, stable, and well organized, and thereby serve as organizers for subsequent steps.

Experiment I

The present experimental studies begin with Piaget's assumption that after a child has advanced to the stage of formal operations, at about age 14 to 16, that he can learn effectively from verbally presented concepts. The purpose of this research is to investigate the effectiveness of using advance organizers of two types, the preview and the model, and their relative effects will be determined in terms of learning, and retaining a system of scientific concepts. The students' attitude towards the imaginary science and the programed instruction, as well as their IQ, science and reading ability will be examined in relation to the measures of learning and retention.

It is hypothesized that the learning and retaining of scientific material can be enhanced in at least three ways: first, by presenting prior to the details of the learning material a preview of the general principles that are to be learned during the presentation of the learning material itself; second, by presenting prior to the presentation of the material itself an analogy or model which serves as a framework into which the details of the learning material can be referred thus enhancing its retention; third, by presenting both a model and a preview prior to the learning material itself, thus providing a direct comparison of the framework of the model with the principles of the science. Since brighter students are probably able to structure materials themselves much more readily than their less gifted companions it is further hypothesized that the facilitation resulting from the use of advance organizers will be more pronounced among the less gifted students.

Method

Learning Material. The imaginary science of Xenograde Systems is a description of a system of satellites which move about a nucleus. The nucleus contains small particles which under certain conditions affect the motion of the satellites. The laws which govern the relationships between these particles in addition to the vocabulary used to describe these laws makes up the subject matter of the science. The Science of Xenograde Systems as used in this study consists of the following sections:

Section H, historical introduction, is a general discussion concerning the use of an imaginary science. This material is not essential to an understanding of the science but is concerned primarily with why one would use an imaginary science. It was written to be similar to the typical first chapter of a text which gives some historical information about the subject, or like the first lecture of a course in which one discusses things about the subject without presenting any details about the subject.

Section M, model, is an analogy which discusses a yoyo machine and suggests that remembering the operation of this machine will assist one in remembering the laws of the science. The yoyo machine makes use of the concepts of a yoyo going up and down a string and the relationship of the size of the yoyo's axle to the speed of rotation. Both of these concepts should already be familiar to students. The science was invented using this model as a guide.

Section D,[2] preview or review, is a summary of the principles and generalizations which are to be learned while studying the science.

---

[2]The letter "D" was used to stand for differentiated since a later section of the program is a problem and used the letter "P."

Section N, nucleonics, is the portion of the science which discusses the nucleus and the particles it contains and the laws which govern their behavior. This section presents all of the details of the science that are concerned with the nucleus and its particles.

Section E, electronics, is the portion of the science which discusses the satellites and describes the laws that govern their movement around the nucleus. This section presents all of the details of the science that are concerned with the satellites and their movements.

Section P, isotopes, presents a hypothetical research problem concerned with the formation of a certain type of Xenograde isotope from a more common variety. The solution of this problem requires application of the laws of nucleonics and electronics as well as deriving some new interrelationships not previously developed.

The sections of the science have been combined as illustrated in Table 1 to form the materials used for the various groups in this study. Group C, conventional, is an attempt to follow the usual approach used in lecture and text to present complex scientific materials, i.e., the subject matter is divided into logical divisions preceeded by a historical introduction and followed by a review. This represents the control condition and does not make use of an advance organizer. Group M presents the model as an example of an advance organizer after which the details of the science are presented. During the presentation of the details of the science (sections N, E, and P) the various aspects of the model relevant to the discussion were called to the learner's attention, e.g., a statement such as: "From the model, remember that a yoyo machine consists of three yoyo's which are attached to a control center that contains small particles called axle

Table 1

Order of Presentation of the Various Sections of the Science of Xenograde Systems to Form the Four Experimental Teaching Programs

Group	Order of Presentation of Sections				
	1	2	3	4	5
Group C	H	N	E	P	D
Group M	M	N	E	P	D
Group P	H	D	N	E	P
Group MP	M	D	N	E	P

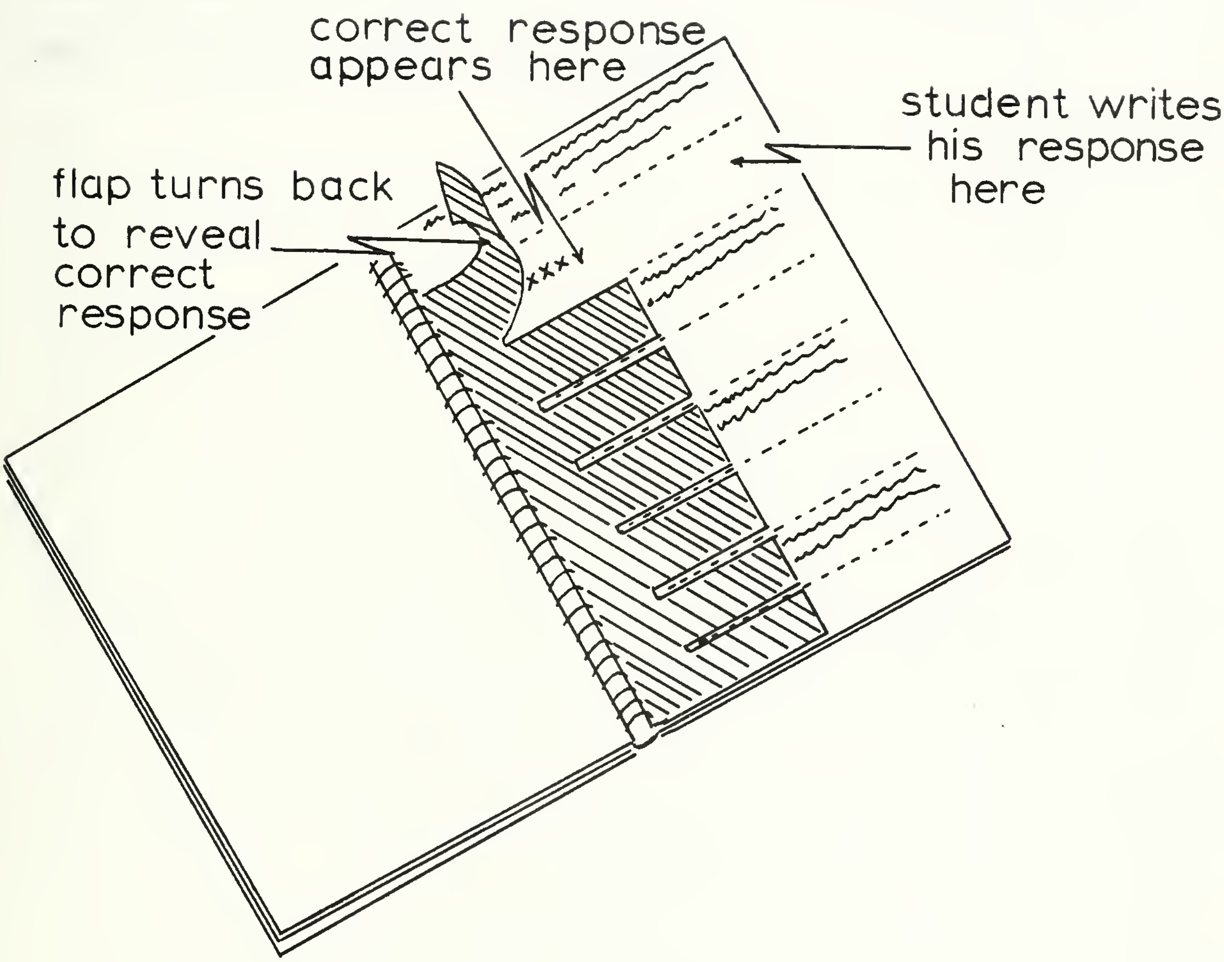
Note. -- The letters stand for historical section (H), model section (M), preview or review section (D), nucleonics section (N), electronics section (E), and problem or isotope section (P). For description of these sections and experimental groups see text.

units.", would be introduced in the program just prior to teaching the student about a Xenograde system consisting of three satellites moving around a nucleus containing small particles called alphons. In groups C and D this statement would be replaced by a statement following the presentation about Xenograde Systems which would review what had just been presented, e.g. "Remember that a Xenograde system consists of three satellites which rotate around a nucleus which contains small particles called alphons." This one-to-one equating was made possible because of the use of programed instruction (see below). Group P presents the preview section prior to the details as an example of an advance organizer. The general principles which are presented in this section are presented one by one during the presentation of the details of the science just prior to the sections dealing with that particular principle. In Groups M and C these statements follow the presentation of the principles and take the form of a review statement. Group MP presents both the model and the preview sections prior to presenting the details of the science. In this case the statements relating the science to the model and presenting the principles are both presented prior to presenting the details. In the other groups, as has already been indicated, these statements take the form of review statements.

Procedure. Programed instruction was chosen as the mode of presentation (see introduction for reasons). In the present instance the only variable in addition to the presentation of the model was order of presentation. As can be seen from Table 1 Groups C and M received the preview section at the end of the learning material in the form of a review, while Groups P and MP were presented this section in the preview position. Throughout the presentation of the details a frame of the program which

presented a principle in preview of things to come in one program appeared in the same form later in the program for the other groups as a review statement. The only exception to this one-to-one equating of the programs was in the case of the model which was paralleled by the historical section. In this case the substance of the material was different, but the number of frames was equal to counter balance any practice effect that the model might have provided.

Many forms of programed instruction exist (see Stolurow, 1961). For the current study a simple linear form of program of the type advocated by Skinner (1958) was adopted. This consists of the presentation of a statement with a missing word or words and requires the student to fill in the missing portion of the statement. By using the principle of vanishing cues the responses require the student to master more and more of the material until he finally has a grasp of the concepts presented (Stolurow, 1961). A form of programed book was used for the presentation of this program. This format is illustrated in Figure 1. The student was required to turn back a cover slip, write his response in the blank space in the next frame on the page and then turn back the next cover slip to check his response. The consecutive frames ran down the page instead of through the book as has been used in the zebra format which several other programs have used (see e.g., Holland and Skinner, 1961). In addition to the programed book each subject was given a booklet containing a number of displays. These consisted of illustrations of the Xenograde Systems, tables of data used in developing the laws of the system, and graphs plotting the relationships of the data. At various places in the presentation of the program the



PROGRAMED BOOK

Fig. 1 -- Programed book used to present the program in Experiment I. Four items appeared on each page. The subject read a step or frame and then wrote his response in the right-hand side of the space following the item. After writing his response he lifted the flap covering the correct response in the left hand side of the space and compared his response with the correct response.

student was requested to look at various displays in this booklet.[3]

Tests. Immediately after learning the program and again after two weeks the subjects were given a test on the Science of Xenograde Systems. These tests consisted of two sections. Section I contained five problems which presented some new data and asked the student to solve the problems by applying the principles which he learned concerning the Science. Section II consisted of 32 four-answer multiple-choice questions. These questions were of two types, Taught Knowledge items, which merely asked the students to recall vocabulary or laws that they learned in the science, and Application items, which presented some data or a problem and asked the students to choose the correct answer from the choices by applying the principles or laws he had learned. These three types of items were used to allow us to compare verbal recall with problem solving and consequently more adequately evaluate degree of comprehension (see introduction). On these items the subjects were asked to respond by crossing off the incorrect answers rather than merely circling the correct answer. The items were then scored according to the scheme illustrated in Table 2. This scheme allowed the subject to receive a score ranging from -3 to +3 on any item. The negative scores compensate for guessing, giving a chance score of 0 for the test. Two forms of the test were constructed. Section I of these two forms contained identical questions but different data. Section II consisted of similar but different multiple choice questions. Approximately half of the subjects were given Form A on the immediate test and then were

[3]A copy of the programed material, the displays, and the test are on file in the University of Illinois Library, Urbana, Illinois in an unpublished volume titled, The Imaginary Science of Xenograde Systems: An Experimental Teaching Program, M. David Merrill, 1964.

Table 2

Scheme for Scoring Multiple
Choice Items on Section Two of the Tests

(Assume "a" is the correct response in
each of the following response patterns)

	Response Patterns						
	I	II	III	IV	V	VI	VII
	~~a~~	~~a~~	~~a~~	a	a	a	a
Response	b	~~b~~	~~b~~	b	~~b~~	~~b~~	~~b~~
Choices	c	c	~~c~~	c	c	~~c~~	~~c~~
	d	d	d	d	d	d	~~d~~
Score	-3	-2	-1	0	+1	+2	+3

Note. -- The letters "a" through "d" represent the four possible answers to each problem. The lines through the letters indicate those answers which a student felt were incorrect. There are other forms that the above patterns can take, e.g., a pattern a ~~b~~ c ~~d~~ also was scored as +2.

given Form B on the two-week retention test. The other half were given Form B on the first test and Form A on the second test.

Questionnaire. Following the second test all subjects were given a questionnaire which requested the following information:

How much have you discussed the Science of Xenograde Systems with your parents and friends? (a) not at all (b) told only my parents (c) told only a few friends (d) discussed it with others who learned it (e) discussed it with friends and others who learned it and tried to clear up places where I didn't understand.

How much did you enjoy the programed book type of instruction? (a) very poor way to learn (b) I didn't enjoy it (c) no opinion one way or the other (d) it's OK (e) very good way to learn.

How much did you enjoy learning the imaginary Science of Xenograde Systems? (a) a complete waste of time (b) I didn't enjoy it (c) no opinion one way or the other (d) it's OK (e) very enjoyable.

Subjects. The subjects were 10th and 11th grade students from a suburban high school. The original sample consisted of 128 students. The students were selected randomly from school records according to the following scheme. Every fifth, sixth, or seventh student on the records in each category was included in the sample, e.g., every fifth 10th grade boy with an IQ above 125 was chosen for a total of sixteen. Every sixth 10th grade girl with an IQ above 125 was chosen for a total of sixteen, etc. until the sample was complete. The sixteen in each category were then assigned to an experimental group from a table of random numbers. Because of absenteeism and failure to complete the second session of learning after completing the first session, or failure to take the test after learning the science, the final sample consisted of 105 subjects distributed as

ing the science, the final sample consisted of 105 subjects distributed as illustrated in Table 3.

The subjects were required to gather in the auditorium for two mornings of regular school days. The total time for most subjects was three hours each day for a total of six hours. Some students had not finished learning the program at the end of the six hours and were allowed to come back in the afternoon of the second day to finish learning the program and to take the test. The second test was administered two weeks later to all subjects.

Results

There was available for each subject Thorndike-Lordge IQ scores, STEP science and STEP reading scores. In addition, questionnaire data were collected as to whether the subject liked programed instruction, whether he liked learning an imaginary science, and how much he discussed the science with his friends between the long term and the immediate tests.

Two-way analyses of variance were run on the IQ, reading and science scores to determine if our process of random sampling had indeed produced groups that were equal on these variables. Tables 4, 5, and 6 indicate the means and the summary of the analysis of variance for each group on each of these variables. As can be seen from this analysis there was a significant difference between IQ levels, but there was no significant difference between the experimental groups. This is the result we would anticipate from our random selection of groups within levels.

The results of a Chi square analysis of the questionnaire data are summarized in Tables 7 through 12. In these tables the number of persons who subscribed to each statement are the observed frequencies. Each

Table 3

Distribution of Final Sample (N = 105)

IQ Level	Groups			
	MP	P	M	C
High	10th boys 3	10th boys 2	10th boys 3	10th boys 5
	11th boys 3	11th boys 3	11th boys 4	11th boys 3
	10th girls 4	10th girls 4	10th girls 2	10th girls 4
	11th girls 3	11th girls 3	11th girls 3	11th girls 3
	n = 13	n = 12	n = 12	n = 15
Low	10th boys 2	10th boys 2	10th boys 4	10th boys 1
	11th boys 3	11th boys 4	11th boys 3	11th boys 3
	10th girls 4	10th girls 4	10th girls 5	10th girls 4
	11th girls 4	11th girls 2	11th girls 5	11th girls 3
	n = 13	n = 12	n = 17	n = 11
Total	n_{mp} = 26	n_p = 24	n_m = 29	n_c = 26

Note. -- Two alternates were assigned to each group to insure adequate sample size. In almost every case, however, more than two were absent thereby reducing the sample as indicated. Where everyone showed up including the alternates the sample is larger than planned. (See group M-low and C-high.)

Table 4

Summary of Analysis on Variance of
Groups on Thorndike-Lordge IQ Scores

Source	df	SS	MS	F
IQ Levels	1	9537	9537	158.4#
Groups	3	85	28.3	.47
Interaction	3	148	49.3	.82
Within	84	5055	60.2	
Total	91	14825		

#$p < .01$

Note. -- Means were as follows: High IQ, group MP, 130; P, 133; M, 132; C, 134; Low IQ, MP, 111; P, 113; M, 114; C, 109.

Table 5

Summary of Analysis of Variance
of Groups on STEP Reading Scores

Source	df	SS	MS	F
IQ Levels	1	8145	8145	57.4#
Groups	3	656	218.7	1.54
Interaction	3	160	53.3	.38
Within	84	11912	141.8	
Total	91	20873		

#p .01

Note. -- Means were as follows: High IQ, group MP, 316; p. 315; M, 313; C, 310; Low IQ, MP, 297; P, 293; M, 299; C, 290.

Table 6

Summary of Analysis of Variance
of Groups on STEP Science Scores

Source	df	SS	MS	F
IQ Levels	1	3195	3195	24.5#
Groups	3	256	85.3	.7
Interaction	3	352	117.3	.9
Within	84	10598	130.4	
Total	91	14401		

#$p < .01$

Note. -- Means were as follows: High IQ, group MP, 292; P, 292; M, 297; C, 292; Low IQ, MP, 280; P, 287; M, 281, C, 280.

Table 7

Chi Square Analysis of Questionnaire Data
for High vs Low IQ Group Concerning Amount of
Discussion Engaged in about the Imaginary
Science Between First and Second Test

IQ Level		Amount of Discussion			
		Little or none			A great deal
		1 & 2	3	4	5
High	E	6.71	12.90	18.58	9.81
	0	7	7	23	11
Low	E	6.29	12.10	17.42	9.19
	0	6	18	13	8

Note. -- Each cell contains the expected (E) and observed (0) frequency of response to the choices listed on the questionnaire. For statements which correspond to items 1 through 5 above see text. $\chi^2 = 8.07^{\#}$, df = 3.

$^{\#}p$ .05.

Table 8

Chi Square Analysis of Questionnaire Data for Four Experimental Groups Concerning Amount of Discussion Engaged in About the Science Between First and Second Test

Group		Amount of Discussion		
		Little or none		A great deal
		1, 2, 3	4	5
MP	E	10.62	10.07	5.31
	0	16	6	4
P	E	8.58	8.13	4.29
	0	9	10	2
M	E	9.81	9.29	4.90
	0	8	9	7
C	E	8.99	8.52	4.49
	0	5	11	6

Note. -- Each cell contains the expected (E) and observed (0) frequency of response to the choices listed on the questionnaire. For statements which correspond to items 1 through 5 above see text. $\chi^2 = 10.60$, df = 6.

Table 9

Chi Square Analysis of Questionnaire Data for High vs Low IQ Group Concerning Degree They Liked Programed Instruction

IQ Level		Opinion of Programed Instruction			
		Disliked			Liked
		1 & 2	3	4	5
High	E	6.13	6.13	15.84	18.9
	0	7	6	16	18
Low	E	5.87	5.87	15.16	18.09
	0	5	6	15	19

Note. -- Each cell contains the expected (E) and observed (0) frequency of response to the choices listed on the questionnaire. For statements which correspond to items 1 through 5 above see text. $\chi^2 = 0.35$, df = 3.

Table 10

Chi Square Analysis of Questionnaire Data for the Four Experimental Groups Concerning Degree They Liked Programed Instruction

Group		Opinion of Programed Instruction	
		Dislike or neutral 1, 2, 3	Liked 4, 5
MP	E	6.97	20.03
	0	5	22
P	E	5.42	15.58
	0	8	13
M	E	6.19	17.8
	0	4	20
C	E	5.42	15.58
	0	7	14

Note. -- Each cell contains the expected (E) and observed (0) frequency of response to the choices listed on the questionnaire. For statements which correspond to items 1 through 5 above see text. $\chi^2 = 4.22$, df = 3.

Table 11

Chi Square Analysis of Questionnaire Data for High vs Low IQ Groups Concerning Degree They Liked Learning an Imaginary Science

IQ Level		Opinion of Imaginary Science				
		Disliked				Liked
		1	2	3	4	5
High	E	2.5	8.8	8.3	20.1	8.3
	0	1	9	9	19	10
Low	E	2.4	8.2	7.7	18.9	7.7
	0	4	8	7	20	6

Note. -- Each cell contains the expected (E) and observed (0) frequency of response to the choices listed on the questionnaire. For statements which correspond to items 1 through 5 above see text. $\chi^2 = 3.23$, df $= 4$.

Table 12

Chi Square Analysis of Questionnaire Data
for the Four Experimental Groups Concerning Degree
They Liked Learning an Imaginary Science

Group		Opinion of Imaginary Science	
		Disliked or neutral 1, 2, 3	Liked 4,5
MP	E	10.62	15.37
	0	11.	15
P	E	8.58	12.41
	0	12	9
M	E	9.80	14.19
	0	8	16
C	E	8.98	13.01
	0	7	15

Note. -- Each cell contains the expected (E) and observed (0) frequency of response to the choices listed on the questionnaire. For statements which correspond to items 1 through 5 see text.
$\chi^2 = 3.63$, df = 3.

subject was allowed to subscribe to only one of the five choices for each question. These results reveal that the high IQ subjects discussed the science outside the experimental situation more than did the low IQ group. However, none of the experimental groups discussed the science outside the experimental situation any more than did the other groups. There was no apparent difference between IQ levels or between experimental groups as to the degree they liked both the programed method of presentation and the content. Seventy-four per cent of the students said they liked programed instruction (subscribed to items 4 and 5), and fifty-seven per cent they liked learning an imaginary science (subscribed to items 4 and 5).

All of the students were tested immediately after learning with a test composed of three types of items, each of which was scored separately. The first and second types were presented in multiple choice format. With the first type, the student was asked to recall some fact or principle that he learned during the presentation of the science. With the second type, the student was asked to apply some principle to date given him in the test item. All of these problems were like those given the students while learning about the science. The third type required the student to apply one or more of the principles of the science to some new data in a situation not previously encountered in learning the science. These three types of items will be called by the names Taught Knowledge, Application, and Problem, respectively.

Tables 13 through 15 report the means and summary of the three way analysis of variance for the 2x2x2 design of the scores on each of the item types on the immediate test. Figure 2 gives the paradigm used for this analysis (Lindquist, 1940). Because of unequal n's in the final sample

Table 13

Summary of Three Way Analysis of Variance on Problem Items for the Immediate Test

Source	df	SS	MS	F
IQ Levels	1	4.80	4.80	3.66
P	1	4.31	4.31	3.29
M	1	1.05	1.05	0.80
IQ x P	1	0.02	0.02	0.01
IQ x M	1	2.02	2.02	1.54
P x M	1	0.04	0.04	0.03
IQ x P x M	1	0.68	0.68	0.52
Within Cells	84	109.78	1.31	
Total	91	122.70		

Note. -- Means were as follows: High IQ, group MP, 1.50; P, 1.83; M, 0.92; C, 1.58; Low IQ, group MP, 1.18; P, 1.27; M, 0.91; C, 0.64. The total possible score for this subtest was 5.

Table 14

Summary of Three Way Analysis of Variance on Application Items for the Immediate Test

Source	df	SS	MS	F
IQ Levels	1	376	376	1.07
P	1	31	31	.09
M	1	39	39	.11
IQ x P	1	192	192	.55
M x P	1	360	360	1.03
IQ x M x P	1	007	007	.02
Within Cells	84	2949	351.1	
Total	91	4057		

Note. -- Means were as follows: High IQ, group MP, 8.67; P, 8.58; M, 6.83; C, 13.60; Low IQ, group MP, 10.18; P, 4.72; M, 1.45; C, 5.09. The total possible score for this subtest was 21.

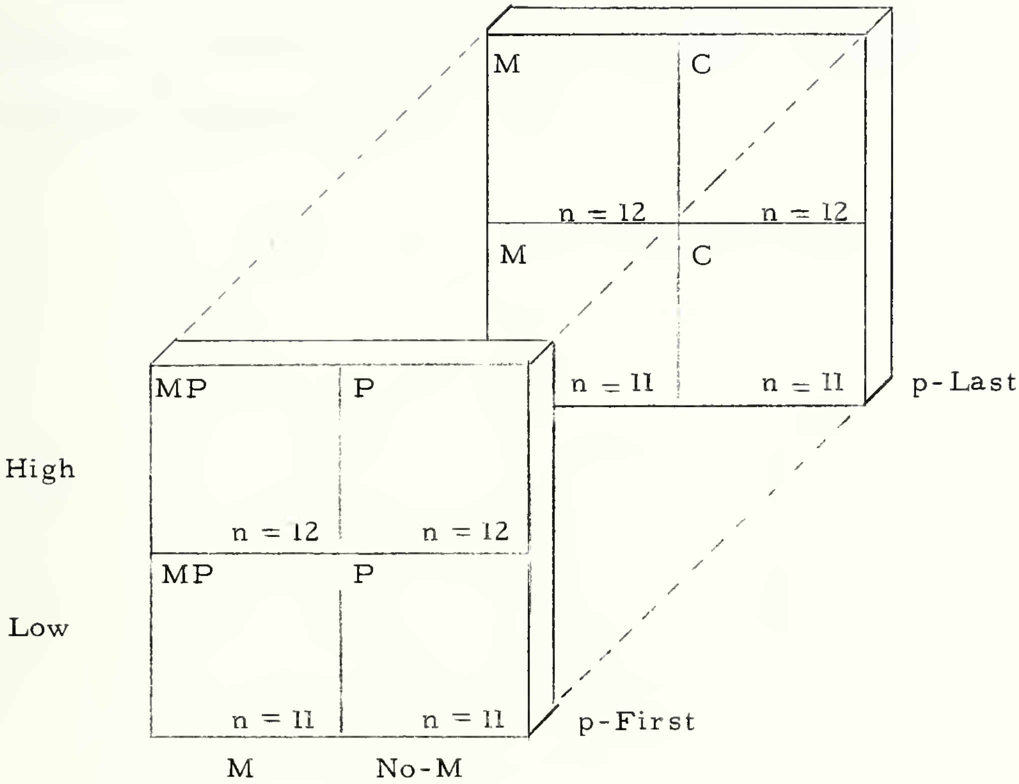
Table 15

Summary of Three Way Analysis of Variance on Taught Knowledge Items for the Immediate Test

Source	df	SS	MS	F
IQ Levels	1	6260	6260	35.60#
P	1	210	210	1.19
M	1	27	27	0.15
IQ x P	1	190	190	1.08
IQ x M	1	217	217	1.23
P x M	1	0	0	0
IQ x P x M	1	102	102	.57
Within cells	84	14,786	176	
Total	91	21,792		

#p .01

Note. -- Means were as follows: High IQ, group MP, 46.0; P, 46.2; M, 44.3; C, 47.9; Low IQ, group MP, 34.2; P, 31.4; M, 29.3; C, 23.5. The total possible score for this subtest was 75.



2x2x2 Design

Fig. 2. 2x2x2 Design Used in Three Way Analysis of Variance of Test Results

High, Low refers to IQ Level; M, No-M refers to model presented First or not presented; p-First, p-Last refers to preview presented first or presented last as review. Total N = 92.

(Table 3) it was decided to discard subjects from some groups. These were chosen randomly and omitted from the 2x2x2 analysis.

The analyses indicate that on the Problem and Application item subtests there are no differences between experimental groups nor between IQ levels. A comparison of the means with the possible score on each subtest reveals that the scores were extremely low. In fact, for the Problem items, over half of the subjects (55) received a score of 0. Only a few subjects scored 3 or 4. For the Taught Knowledge subtest there was no difference between experimental groups, but there was a significant difference between IQ levels. It is interesting to note, however, that the means for the lower IQ level are in the predicted order, but because of large variance and small n's this trend is not significant and could have resulted from chance. Failure to get significant differences between levels for the other subtests probably resulted from the fact that no one was able to solve very many of these problems and this was equally true of high IQ subjects and low IQ subjects.

Tables 16 through 18 report the group means and a summary of the three way analyses of variance for the 2x2x2 design for each set of scores on the two-week retention test. The analyses indicate that for all three types of items, the high IQ students remember more than the lower IQ students. However, the differences among the four experimental groups were not significant. Two interactions were significant. The first is between IQ level and the presence or absence of the model during learning. It indicates that high IQ students are aided in retaining principles necessary to solve application problems if they first learn a model, but that low IQ students are better able to retain principles necessary for solving these

Table 16

Summary of Three Way Analysis of Variance
on Problem Items for the Retention Test

Source	df	SS	MS	F
IQ Levels	1	10.49	10.49	11.69#
P	1	0.04	0.04	.04
M	1	0.04	0.04	.04
IQ x P	1	0.18	0.18	.20
IQ x M	1	0.01	0.01	.01
M x P	1	1.57	1.57	1.75
IQ x M x P	1	0.03	0.03	.03
Within Cells	84	75.38	0.897	
Total	91	87.74		

#p .01

Note. -- Means were as follows: High IQ, group MP, 0.92; P, 1.08; M, 1.25; C, 1.00; Low IQ, group MP, 0.27; P, 0.55; M, 0.55; C, 0.18. The total possible score for this subtest was 5.

Table 17

Summary of Three Way Analysis of Variance
on Application Items for the Retention Test

Source	df	SS	MS	F
IQ Levels	1	316	316	14.70$^{\#}$
P	1	23	23	1.06
M	1	11	11	.51
IQ x P	1	0	0	0
IQ x M	1	112	112	5.21$^{\#\#}$
M x P	1	0	0	0
IQ x M x P	1	18	18	.84
Within Cells	84	1804	21.5	
Total	91	2284		

$^{\#}p$ .01

$^{\#\#}p$ .05

Note. -- Means were as follows: High IQ, group MP, 8.25; P, 6.08; M, 6.58; C, 5.92; Low IQ, group MP, 1.55; P, 5.54; M, 1.45; C, 3.45. The total possible score for this subtest was 21.

Note. -- The combined means showing the interaction are as follows: High IQ-model presented (n = 24), 7.4; -no model presented, 6.0; Low IQ-model presented (n = 22), 1.5; -no model presented, 4.5.

Table 18

Summary of Three Way Analysis of Variance on Taught Knowledge Items for the Retention Test

Source	df	SS	MS	F
IQ Level	1	7033	7033	38.9#
P	1	10	10	.06
M	1	663	663	3.66
IQ x P	1	384	384	2.12
IQ x M	1	48	48	.21
M x P	1	835	835	4.61##
IQ x M x P	1	158	158	.87
Within Cells	84	15,224	181	
Total	91	24,355		

#$p < .01$

##$p < .05$

Note. -- Means were as follows: High IQ, group MP, 33.7; P, 34.2; M, 34.8; C, 42.3; Low IQ, group MP, 21.5; P, 19.5; M, 09.1; C, 24.7. The total possible score for this subtest was 75.

Note. -- The combined means showing the interaction are as follows: Model presented-preview (n = 23), 27.8; Model presented-review (n = 23), 22.5; No model-preview (n = 23), 27.2; No model-review (n = 23), 33.9.

problems if they do not have the model. The second significant interaction indicates that for maximum retention of taught knowledge about the science neither a model nor a preview should be used; retention is next best when a preview is used whether or not a model also is presented with it; and retention is poorest if no preview is used but a model is.

Intercorrelations. Table 19 gives the results of correlating the following variables: sex, grade in school, IQ, STEP science scores, STEP reading scores, Test sequence (A or B presented first), the three questions from the questionnaire (i.e., Q_1 How much did you enjoy the programed book type of instruction? Q_2 How much did you enjoy learning the imaginary Science of Xenograde Systems? Q_3 How much have you discussed the Science of Xenograde Systems with your parents and friends?), and each of the test scores (i.e. Immediate test -- Problem score, Application score, and Knowledge score; and retention test -- Problem score, Application score, and Knowledge score.) From this table it can be seen that IQ correlates slightly with reading scores but that the correlation coefficient between IQ and Science scores fails to reach significance at the .05 level. IQ is also correlated with each of the test scores being more highly correlated with the knowledge items on both immediate and retention tests and correlated least with problem items on both tests. Science scores correlated highly with reading scores but all of the coefficients of science scores with test scores failed to reach significance. Reading scores correlated slightly with the amount the subject discussed the science with others, but failed to correlate significantly with the test scores.

Table 19

Correlation Matrix for the Variables Considered

	Sex	Grd	IQ	Sci	Read	Seq.	Q_1	Q_2	Q_3	T_1P	T_1A	T_1K	T_2P	T_2A
Grade	-085													
IQ	-086	-037												
Sci.	046	070	170											
Read.	082	012	244#	696#										
Seq.	-078	-067	053	-072	036									
Q_1	070	025	166	131	170	-033								
Q_2	135	-008	131	102	181	005	751#							
Q_3	132	-091	188	157	256#	003	739#	715#						
T_1P	-006	041	327#	067	076	-041	192	173	110					
T_1A	-073	028	367#	-007	029	-323#	145	136	080	400#				
T_1K	-069	002	580#	155	185	-178	075	120	090	444#	593#			
T_2P	-123	085	318#	093	135	064	090	038	036	348#	285#	458#		
T_2A	-051	021	403#	002	152	013	124	190	031	424#	307#	526#	452#	
T_2K	-087	-102	595#	051	100	073	125	125	101	370#	486#	650#	418#	558#

Note. -- Seq. indicates whether the subject had form A for the immediate test and form B for the retention test or the reverse. Q_1 was, How much did you enjoy the programed book type of instruction? Q_2 was, How much did you enjoy learning the imaginary Science of Xenograde Systems? Q_3 was, How much have you discussed the Science of Xenograde Systems with your parents and friends? T_1problem, T_1application, T_1knowledge are scores on the immediate test for each type of item. T_2problem, T_2application, and T_2knowledge are scores on the retention test for each type of item.

#p .01

Test sequence has a significant effect. If test Form A is given before Form B, then negative correlation with Application items results on the immediate test. This suggests that the Application items on Form B were more difficult than the application items on Form A (Form A was coded 0 and B coded 1 for the correlation). Since there were not equal numbers of subjects receiving each form in each group this difference in difficulty would affect the means reported in Table 14. A separate analysis of variance for each form was run to check the possibility that this difference in difficulty might alter the results reported. It was found that there was still no statistical difference between the experimental groups on either form of the test. One might question the reason for this correlation on the immediate test and not on the retention test. The author's hypothesis is that on the retention test the easier Application items on Form A were also missed thereby equalizing the scores with those of Form B which were missed the second time as well as the first time and thus lowering this correlation between form sequence and Application items to 0.

The correlations for the questionnaire data indicate that if a student liked learning by means of programed instruction he also liked to learn an imaginary science and that he also discussed the science more outside the experimental situation. What this probably means is that if a student liked participating in the study he liked everything about it and discussed it with his friends and family. If, on the other hand, he didn't like participating he didn't like anything about it and didn't discuss it with anyone much.

There is a positive correlation between all of the test scores. Some interesting patterns of relationships can be observed from Table 19. First, the correlations between part scores on the same test, while positive,

are low enough to indicate that the various types of items probably measured somewhat different abilities. On both the retention and immediate achievement tests, both the Application and the Taught Knowledge items are more highly correlated with each other than are the Problem items with either the Application or the Taught Knowledge items. Second, a high Taught Knowledge score on the immediate test was more likely to be followed by a high score on the retention test than was a high Application or Problem score. This indicates that Taught Knowledge information was remembered, but was not sufficiently understood to solve problems.

Table 20 reports multiple correlation coefficients for each test with the other nine variables. The same relationship as reported above can be seen here, i.e., that the variables are more useful in predicting Knowledge subtest scores than they are in predicting Problem subtest scores, Application subtest scores are predicted at a level falling somewhere in-between these two predictions. In all cases, however, there is a great deal (54 to 79%) of unaccounted for variance in the predicted scores.

Table 20

Multiple Correlation Using
Sex, Grade, IQ, Science, Reading, Test Form,
and Questionnaire Data to Predict Test Scores

T_1P	T_1A	T_1K	T_2P	T_2A	T_2K
456#	539#	677#	487#	527#	651#

p .01

Note. -- T_1P, T_1A, T_1K, T_2P, T_2A, and T_2K represent Problem (P), Application (A), and Knowledge (K) items for the immediate (T_1) and the retention (T_2) tests.

Experiment II

This study compared two of the four conditions used in Experiment I.

Method

Subjects. Twenty two college juniors and seniors who were enrolled in a summer session of Educational Psychology learned the program as part of their course requirements.

Learning Material. The same material was used as for groups MP and C in Experiment I.

Procedure. The pages of the program books used for the high school students were in a format that allowed them, if unbound, to be used in Min/max teaching machines manufactured by TMI-Grolier. Operation is as follows: Typewriter-like rollers automatically present the programed learning material (printed on $8\frac{1}{2}$ X 11 inch paper) in order. The student views each frame in the program through a tilted window. Below the viewing space is a cut-out section where the student can write his response. On the left side of this cut-out section is a masked section which conceals the correct answer. By turning the sheet forward the student's response and the correct response appear under the plastic viewing window. A gear arrangement prevents the student from turning the program backward and hence prevents him from changing his response after seeing the correct response.

Students signed up for a time and came into the laboratory according to a schedule. They were allowed to work for two hours at a sitting and were asked to return for as many sessions as necessary to learn the program.

Using a table of random numbers the subjects were divided into two groups. Ten subjects learned with the program presenting the model and

the preview section -- the same as group MP in Experiment I. Twelve subjects learned the program without the model and with the review rather than the preview -- the same as group C in Experiment I.

Tests. Forms A and B of the tests given to the high school subjects were combined into a single test and given to the college students. There was therefore only one test given immediately following the learning of the program. The subjects were instructed to circle the correct choice on the multiple choice items rather than to mark off incorrect responses as in Experiment I. Each test item was scored as correct or incorrect and the total score was merely the number of correst items.

Results

In Experiment II the group means for the Taught Knowledge subtest were as follows: group MP, 33.5; group C, 35.5. The total possible score on the Taught Knowledge subtest was 50. A comparison of the two groups on these scores yielded a t of 0.91 which is not significant. This indicates that there was not a significant difference in amount learned between groups.

The means for the Application items were 6.60 for group MP and 7.92 for group C. The total possible score on Application items was 14. A comparison of the two groups on these scores yielded a t of 1.13 which is not significant. The groups, therefore, did not differ on the Application subtest. It should also be pointed out that there is no correction for guessing for Taught Knowledge or Application scores and it is likely that the means reported were inflated slightly by guessing.

The means for the Problem subtest were 3.4 for group C and 2.4 for group MP. The total possible score is 10. Since these were completion

items there is likely no inflation from guessing. A comparison of these means yielded a *t* of 1.09 which is not significant. Again there is no significant difference between groups.

The average time necessary to complete the program was four hours and 37 minutes. This average was exactly the same for both groups indicating that there is no time advantage for one program over the other.

Discussion

The results reported do not support the hypothesis posed at the beginning of this paper. One might conclude that when using a small step-size, linear program to teach complex scientific materials, presenting either a model or preview or both before presenting the details of the science does not facilitate either the learning or the retaining of the material.

Two significant interactions were found which require some qualification of the above conclusion. The first interaction effect showed that high IQ students were aided in the retention of principles necessary for solving Application items by learning a model while the less gifted were aided by getting a review without the model. Clearly this is contrary to the hypothesis that advance organizers facilitate retention more among the less gifted than the gifted. Perhaps this means that the level of learning was so low in this study that only the gifted were able to understand and therefore remember the principles necessary for solving the application items and hence were aided by the model in retaining these principles while the less gifted were unable to understand or learn them in the first place. The fact that the same interaction was not obtained for the problem items probably indicates that even the brighter students were unable to understand

or learn the principles necessary for solving these problems and consequently were unaided by knowledge of the model.

The second interaction shows that retention of the Taught Knowledge items is best when neither a model nor a preview is presented in the program. This seems to indicate that students remember more when less is presented for them to remember but that a preview can aid retention when a model also is presented. The big question, in the light of Murphy's (1962) results, is why was the no-model - review condition (the summary reviewed) better than the no-model - preview condition? According to Murphy's findings we would predict the opposite result.

Stolurow (1960; 1961) has suggested that when programed instruction is truly effective the correlation between IQ and performance typically drops to zero. From Table 19 it can be seen that IQ correlates higher than any of the other measures obtained with performance. According to this criteria then, the current program was not very effective. The following discussion will attempt to analyze the causes of its ineffectiveness.

Appendix A contains a summary of an item analysis for the two tests. The analysis used was the one suggested by Davis (1946) which is based on the upper and lower 27 per cent of the subjects. Total test scores (Taught Knowledge and Application items scores combined) were used to divide the students into these groups. In addition to Davis' difficulty index and discrimination index, this table also indicates the percentage of students who passed each item. Separate indexes and percentages are given for the high school and college students (Experiments I and II). The second column of the table classifies each item as to the

type of information required to pass the item. Column three indicates[4] the original classification of the item into Problem, Application, or Taught Knowledge types as was explained in the procedure section of the first experiment.[5] The last three columns of the table indicate whether the item is one of the hardest or easiest 25 per cent of the items and whether it was hard or easy for one or both groups of subjects.

Table 21 is a summary of the information contained in the last three columns of the item analysis in Appendix A. This table indicates how many items were of each type and how many of each type of item were hardest and easiest for the subjects. Examination of this table indicates that the easiest items were those calling for knowledge of terminology and specific facts and that the hardest items were application and knowledge of principles and generalizations, with six knowledge of specific fact items and only one knowledge of terminology item. Closer examination of the particular items involved indicates that of the two application items (A-15 and A-29) which

---

[4]This classification was made by Carl Bereiter, who originally conceived of and invented the Science of Xenograde Systems. The categories used were some of those suggested by Bloom (1956). Bereiter was asked to make this logical classification because of his acquaintance with the science and because at the time he made the classification he was unaware of any data concerning which items were easiest, hardest, etc.

[5]It should be noted that there is a correspondence between these two independent classifications. In the Bloom classification the problem items and all but five of the application items from the original classification are called application. All of the taught knowledge items are distributed in the other three Bloom categories according to knowledge of what. The data were analyzed as reported for the following reasons: Problem items and application items (original classification) were scored differently and therefore not easily combined for analysis as would be necessary if Bloom classification served as the basis. The various Bloom knowledge types were not distributed evenly in the two forms and made comparison using both forms difficult and separate comparison impractical because of the small n's that would result. All of the knowledge items were therefore grouped together in the Taught Knowledge subtest.

Table 21

Number of Items of Each Type Falling in the Easiest and Hardest 25 Per Cent of 74 Items

Item Difficulty Group	Item Types: Knowledge of Terminology	Knowledge of Specific Facts	Knowledge of Principles and Generalizations	Application	Total
Easiest 25% one or both groups	12	8	0	2	22
Middle 50% both groups	3	16	5	4	28
Hardest 25% one or both groups	1	6	5	12	24
Total number of items in each class	16	30	10	18	74

Note. -- No item was easiest for one sample (high school or college) and hardest for the other. If it was in an extreme for only one group it was always in the middle for the other group.

were among the easiest items, neither was in the easiest 25 per cent for both groups. For the high school group item A-15 was in the easiest 25 per cent. This item asks for the migration rate of some data presented. Technically it is application but unlike most of the other application items it requires little more than knowledge of the term "migration rate." Item A-29, which was in the easiest 25 per cent for the college students, requires a simple substitution of values from a table into a formula which was well learned by all subjects (see item A-16). For the most part the remainder of the application items required application of two or more principles in order to solve the problem. It is interesting to note that items A-3 (Problem) and B-3 (Problem) ask for exactly the same behavior as item A-29 except that in this case the values needed had to be extracted from a complete set of data for an entire Xenograde System. It will be recalled that over half (55) of the subjects in Experiment I received a score of 0 on the problem items. Examination of the table in Appendix A indicates that only a small percentage of the subjects passed any of these items -- items 2 and 3 being the easiest.

Based on the above analysis it is observed that the current program is quite effective in teaching terminology and some specific facts. On the other hand it is not very effective in teaching the application of principles and generalizations to the solution of problems. Several factors may account for this. First, the results of a reading level analysis using the Farr-Jenkins-Paterson method (1951) shows that the vocabulary of the program is at a college or late high school level. The Dale-Chall method (1948) places the average reading level at grades 13 to 15. Based on these analyses it can be reasonably assumed that the material was difficult

reading for most subjects. Note, however, that there is not a significant correlation between STEP reading scores and test scores. This seems to indicate that something other than reading difficulty contributed to the lack of understanding.

The author feels that when one uses a linear small step program such as the one used in this study, there is a tendency to overteach vocabulary and verbal statements but a failure to teach application and understanding. When an author must find a response for each frame it is a natural tendency to use the new vocabulary for the response. This gives the student ample opportunity to practice the new words and hence to overlearn them. However, because of the small step size and the prompting techniques used the student has little opportunity to solve a multistage problem or to apply more than one principle at a time to the solution of a problem. He therefore does not have the opportunity to learn application and problem solving techniques. In order to teach these behaviors this author feels that one must give the student some practice in applying the principles of the science.

Gagne and Brown (1961) found that students who were given sequential practice in problem solution and were forced to recall necessary principles while solving these problems (their guided discovery group) were able to transfer to new problems more successfully than students who merely had practice in applying a principle to find values but were not forced to recall the principles themselves (their rule and example group). Groups left on their own, without guidance in recall (their discovery group) fell between the other two extremes. Examination of some frames in the current program (see frames E-100 to E-139) will indicate that the procedure used was

similar to Gagne and Brown's rule and example group. Perhaps changing the presentation to allow more guided practice in problem solution would increase the level of learning for these materials.

It will be recalled that one of the reasons given for using programed instruction as the medium of presentation was that it allowed monitoring the students responses and controlling the degree of mastery for any given sequence of items (see introduction). It will be noted, however, that such monitoring presupposes a branching type program. In preparing materials for future testing of this hypothesis the author also feels that, in addition to increased step size and guided practice in problem solving, the material must be presented in such a way that it allows monitoring thereby truely controlling mastery of any given sequence by testing the students before allowing them to proceed. This means that it will be necessary to branch them through the same or supplementary material when they fail to reach a specified criterion of performance on each test. By requiring problem solving behavior on these monitoring tests one could also assure the development of the behavior desired.

A review of the program by the author also prompts the conclusion that as written it tends to stress a rote learning process or memorization rather than understanding. That is, instead of adequately explaining a relationship to the student a principle is presented and the student left to memorize the statement. In this way the program hinders understanding that might easily have been achieved were a more explanatory approach used to present the material. For example, frames E-100 to E-128 present Carl's Law as a formula to learn. Were one to point out the relationship involved on a graph before presenting the mathematical formulation the

student would likely see the relationship more clearly.

It will be recalled that Ausubel (1962a, 1963) felt that subsumption theory applies only to a meaningful verbal learning process as opposed to rote learning (see introduction). While the learning material itself (i.e. the imaginary science) is potentially meaningful (i.e nonarbitrary relatability to relevant concepts in cognitive structure) it may have been presented in such a way that it lacked meaning (i.e. failed to be related to the particular structure of some of the students) for many of the students. For one to adequately conclude, as we tentatively did above, that a model and/or a preview does not facilitate learning and retaining complex materials, it would be necessary for the material to be learned in a more meaningful way by the students, that is, one would want better evidence that the students understood the material as measured by their ability to apply principles to the solution of problems. Since very few of the subjects in either experiment were able to apply the principles and generalizations to the solution of simple problems it might be inferred that many of the students adopted a set to memorize the material and that the learning which did take place was probably largely rote in nature. One would be justified in not rejecting the null hypothesis of no difference only if one obtained more convincing evidence that the subject matter was meaningfully learned by the students.

Summary

Based on Ausubel's Subsumption theory and its implications for the use of advance organizers, it is hypothesized that presenting a model and/or a preview prior to the presentation of complex verbal materials will facilitate the learning and retention of those materials.

In Experiment I four groups of high school students were divided into high and low IQ groups. Group I was presented a model and a preview prior to learning a complex imaginary science; Group II was presented the model prior to the science and received a review in place of a preview; Group III was presented a preview prior to the science; and Group IV was presented only a review but no model or preview. The same mode of presentation was used for all groups -- linear programed booklets. Students were tested both immediately after completing the program and two weeks later.

In Experiment II two groups of college students learned the materials that were learned by Groups I and IV in Experiment I. The mode of instruction was by Min/max teaching machine. The subjects were tested immediately following learning.

Results indicated no significant main effects; there were two significant interaction effects: (a) Retention as measured by Application items was best for high IQ students when presented a model but best for less gifted students when no model was presented. (b) Retention as measured by items measuring Taught Knowledge was best when no model or preview was presented and poorest when only a model but no preview was presented.

An analysis of test performance seems to indicate that the teaching machine program was effective in teaching knowledge of terminology and knowledge of specific facts but was ineffective in teaching understanding necessary for problem solving. It was suggested that before one would be justified in failing to reject the null hypothesis one would want to replicate the experiment with a revised program that would enable students to attain a higher level of understanding as measured by problem solving ability.

References

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. J. educ. Psychol., 1960, 51, 267-272.

Ausubel, D. P. In defense of verbal learning. Educ. Theory, 1961, 11, 15-25.

Ausubel, D. P. A subsumption theory of meaningful verbal learning and retention. J. gen. Psychol., 1962, 66, 213-224. (a)

Ausubel, D. P. A transfer of training approach to improving the functional retention of medical knowledge. J. Med. Educ., 1962, 37, 647-655. (b)

Ausubel, D. P. Congitive structure and the facilitation of meaningful verbal learning. Paper read at Amer. Psychol. Ass., St. Louis, Sept., 1962. (c)

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune and Stratton, 1963.

Ausubel, D. P. & Fitzerald, D. Organizer, general background, and antecedent learning variables in sequential verbal learning. J. educ. Psychol., 1962, 53, 243-249.

Bloom, B. S. (Ed.) Taxonomy of educational objectives: the classification of educational goals. New York: David McKay Co., 1956.

Brownell, W. A. & Simms, V. M. The nature of understanding. Yearb. nat. Soc. Stud. Educ., 1950, 49.

Dale, E. & Chall, Jeanne, S. A formula for predicting readability: instructions. Educ. Res. Bull., 1948, 27, 37-54.

Davis, F. B. Item-analysis data: their computation, interpretation, and use in test construction. Harvard Educ. Papers, 1946, No. 2.

Farr, J. N., Jenkins, J. J. & Paterson, D. G. Simplification of Flesch reading ease formula. J. appl. Psychol., 1951, 35, 333-337.

Gagne, R. M. & Brown, L. T. Some factors in the programming of conceptual learning. J. exp. Psychol., 1961, 62, 313-321.

Gagne, R. M. & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychol. Monogr., 1961, No. 518.

Holland, J. G. & Skinner, B. F. The analysis of behavior. New York: McGraw Hill, 1961.

Hunt, J. McV. Intelligence and experience. New York: Ronald Press, 1961.

Inhelder, Bärbel & Piaget, J. The growth of logical thinking from childhood to adolescence: an essay on the construction of formal operational structures. Trans. by Anne Parsons & S. Milgram. New York: Basic Books, 1958.

Lindquist, E.F. Statistical analysis in educational research. Boston: Houghton Mifflin, 1940.

Murphy, H.D. The effect of different locations of chapter summary on immediate and delayed recall. Unpublished doctoral dissertation, Peabody College, 1962.

Robinson, Francis P. Effective study. New York: Harper, 1946.

Skinner, B.F. Teaching machines. Science, 1958, 128, 969-977.

Stolurow, L.M. Teaching machines and special education. Educ. and Psychol. Msmt., 1960, 20, 429-448.

Stolurow, L.M. Teaching by machine. Cooperative Res. Mongr., 1961, No. 6, Washington D.C.: Gov. Printing Office.

Appendix A#

Item Analysis for Test Items for
High School and College Groups

Item Identification			High School Immediate test			College combined tests			25% hardest & easiest		
Item number	Classification##	Item type	% passing	Difficulty index	Discrimination index	% passing	Difficulty index	Discrimination index	Both groups	High School	College
Form A -- Section I											
1	A	P	18	35	17	23	18	28	H		
2	A	P	41	47	20	45	30	47			
3	A	P	32	37	57	41	38	58		H	
4	A	P	11	24	36	9	1	0	H		
5	A	P	25	37	10	32	38	58	H		
Form B -- Section I											
1	A	P	6	14	22	5	1	0	H		
2	A	P	33	44	9	45	18	28		H	
3	A	P	51	56	46	59	53	81			
4	A	P	16	35	53	23	38	58	H		
5	A	P	6	21	32	14	18	28	H		
Form A -- Section II											
1	KSF	R	73	59	18	73	53	81			
2	A	A	47	45	50	41	30	47			
3	KSF	R	65	61	26	68	41	32			

#Five pages to this table.

Appendix A

Item number	Classification##	Item type	% passing	Difficulty index	Discrimination index	% passing	Difficulty index	Discrimination index	Both groups	High School	College
4	KSF	R	39	48	47	45	38	58			
5	KT	R	89	76	36	95	75	-37	E		
6	KT	R	91	76	36	95	75	37	E		
7	KP&G	R	74	61	39	64	66	52			
8	A	A	35	45	38	36	34	19	H		
9	KP&G	R	53	50	42	73	47	48			
10	KSF	R	77	63	36	55	38	58		E	
11	KT	R	97	88	19	100	99	0	E		
12	KSF	R	95	76	36	100	99	0	E		
13	KP&G	R	39	48	36	32	38	58			H
14	KP&G	A	75	63	23	82	59	64			
15	A	A	89	67	50	77	53	81		E	
16	KSF	R	89	73	41	95	75	37	E		
17	KT	R	86	66	17	95	75	37	E		
18	KSF	R	16	35	17	18	1	0	H		
19	KT	R	49	51	30	50	47	14			
20	KT	R	82	63	56	86	66	0	E		
21	KSF	R	79	61	39	86	66	52	E		
22	KSF	R	49	54	13	18	18	28			H
23	KT	R	91	72	11	86	75	37	E		

Appendix A

Item number	Classification##	Item type	% passing	Difficulty index	Discrimination index	% passing	Difficulty index	Discrimination index	Both groups	High School	College
24	KSF	R	86	67	50	82	59	64		E	
25	KSF	R	70	58	48	73	59	15			
26	KP&G	A	60	54	43	73	53	81			
27	KSF	R	25	43	33	14	1	0	H		
28	KSF	R	82	67	29	100	99	0	E		
29	A	A	70	59	64	91	59	64			E
30	A	A	47	51	30	36	38	58			H
31	KSF	R	44	51	22	68	59	15			
32	KSF	R	32	45	8	50	41	- 32		H	
Form B -- Section II											
1	A	A	35	37	32	27	25	0	H		
2	KT	R	35	40	20	23	38	-58	H		
3	A	A	29	42	42	27	30	47	H		
4	A	A	60	60	62	73	53	0			
5	KT	R	69	59	16	68	50	100			
6	KSF	R	46	50	67	55	47	14			
7	KSF	R	60	60	62	82	66	0			
8	KSF	R	58	59	29	50	47	14			
9	KSF	R	67	60	37	55	53	29			

Appendix A

Item number	Classification##	Item type	% passing	Difficulty index	Discrimination index	% passing	Difficulty index	Discrimination index	Both groups	High School	College
10	KP&G	A	35	42	14	32	30	47	H		
11	KSF	R	65	36	3	82	66	52			
12	KP&G	A	33	42	27	59	53	0		H	
13	KSF	R	42	40	22	23	18	28			H
14	KSF	R	40	46	37	64	47	48			
15	KT	R	81	63	6	68	53	29		E	
16	KT	R	71	65	53	91	66	52			E
17	KSF	R	48	50	- 5	32	38	58			H
18	KT	R	82	68	43	95	99	0	E		
19	KT	R	100	99	0	100	99	0	E		
20	KP&G	R	25	32	- 7	45	41	32		H	
21	KSF	R	81	68	22	95	75	- 37	E		
22	KSF	R	52	48	19	73	44	72			
23	KT	R	75	59	29	100	99	0			E
24	KP&G	A	75	68	49	82	59	64			
25	KP&G	A	25	40	22	32	18	28	H		
26	KSF	R	50	59	29	82	66	52			
27	KT	R	65	59	29	68	53	29			
28	KSF	R	65	54	27	73	75	37			

Appendix A

Item number	Classification##	Item type	% passing	Difficulty index	Discrimination index	% passing	Difficulty index	Discrimination index	Both groups	High School	College
29	KT	R	83	68	49	91	99	0	E		
30	KSF	R	48	56	46	68	53	81			
31	KSF	R	75	68	49	100	99	0			E
32	KSF	R	63	54	51	73	59	64			

##A means application, KT means knowledge of terminology, KSF means knowledge of specific fact, and KP&G means knowledge of principles and generalizations.

Note. -- The last three columns indicate which items were the hardest or easiest items for one or both groups. Since there were 74 items the top 19 and bottom 18 items in terms of per cent passing are indicated by the symbols H (hard) and E (easy). This represents approximately the top and bottom 25 per cent of the items. In terms of per cent passing the hard items were passed by 38% or less in the high school group and by 40% or less in the college group. The easiest items were passed by 80% or more in the high school group and by 85% or more in the college group.

I 2 6 p

TRAINING RESEARCH LABORATORY

Department of Psychology Bureau of Educational Research
University of Illinois 8 Lincoln Hall Urbana, Illinois

THE EFFECTS OF SEQUENCE AND STRUCTURE ON COMPLEX CONCEPT FORMATION

Daniel J. Davis

TECHNICAL REPORT NO. 4

Psychological and Educational Factors in
Transfer of Training
Phase I

January, 1964

U. S. Office of Education
Contract 2-20-003

Lawrence M. Stolurow
Principal Investigator

TRAINING RESEARCH LABORATORY
Bureau of Educational Research
Department of Psychology
University of Illinois
Urbana, Illinois

THE EFFECTS OF SEQUENCE AND STRUCTURE ON COMPLEX CONCEPT FORMATION

Daniel J. Davis

TECHNICAL REPORT NO. 4

January, 1964

Psychological and Educational Factors in Transfer of Training
Phase I

Principal Investigator:
Lawrence M. Stolurow

Project Sponsor:
Educational Media Branch
U. S. Office of Education
Title VII
Contract No. 2-20-003

The Effects of Sequence and Structure on Complex Concept Formation

Daniel Davis
University of Illinois

BACKGROUND

Consider a concept formation task in which there are two cues that are relevant to the solution and two cues that are not. Also, of the two cues that are relevant, one is more relevant than the other. That is, solutions based only on the more relevant cue will be closer to the correct solution than those based only on the less relevant one.

For such a task it has been found that subjects can learn the relevance of stimulus parameters simultaneously. Through informational feedback with knowledge of results they learn to ignore the irrelevant cues and to correctly weight and combine the relevant cues into a complex concept (Azuma, 1960, McHale and Stolurow, 1962).

For a similar task in which there was one relevant cue and two irrelevant cues, Detambel and Stolurow (1956) showed that sequencing is an important factor in the effectiveness of training. In particular, great improvement results when the following conditions are met:

a. When the value of the relevant cue changes on adjacent trials, the values of the irrelevant cues remain fixed.

b. When the value of one or both of the irrelevant cues changes, the value of the relevant cue remains fixed.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] relevant [illegible] relevant [illegible] more relevant than the [illegible] only on the [illegible] overlap [illegible] the lower [illegible].

[illegible] relevance [illegible] informational feedback with [illegible] of trials [illegible] irrelevant cues [illegible] only [illegible] combine [illegible] concept [illegible], [illegible] Shepard [illegible].

[illegible] similar [illegible] relevant [illegible] (1956) showed that [illegible] important [illegible] particular [illegible] when the following conditions [illegible].

1. [illegible]

2. [illegible] relevant [illegible] given.

The above conditions determined what they called "asynchronous trials" as compared to "synchronous trials" in which all cues were free to vary simultaneously.

In this study several ways of structuring and sequencing the early trials of a complex concept formation task were compared. The comparison was made on the basis of transfer to later trials which were completely unstructured or synchronous.

The training trials were divided into two main segments:

1. Asynchronous segment (A)--one relevant and one irrelevant cue were held constant while one relevant and one irrelevant cue were free to vary.

2. Synchronous segment (S)--all cues were free to vary.

The asynchronous segment was divided into two parts:

1. A--MAX.--The more relevant cue was free to vary.

2. A--MIN.--The less relevant cue was free to vary.

The four possible orders of presenting the above conditions were compared with each other and with a control condition in which only synchronous trials were given during training. The purpose of this was to obtain answers to the following questions:

1. What is the effect of adding the asynchronous trials?

Based on the study of Detambel and Stolurow it is expected that during training the asynchronous groups should do better, but it is not at all certain that they will transfer to the synchronous situation. Whether they do or not

would seem to depend on how they form the concept. If they operate on each cue independently and then combine them, there should be positive transfer. The reason for this is that on asynchronous trials they can direct all their attention to one cue at a time and later try to combine them. If, on the other hand, they use relationships between the cues, the asynchronous trials should be of no help.

2. What is the best order of presentation of the asynchronous (A) and synchronous (S) training trials?

It is expected that the order S-A is better than A-S. In the former case, the subject is familiarized with the situation to which he must eventually transfer. Therefore, on the A-trials he has a reference on which to base his hypotheses.

3. During the asynchronous trials is it better to present the more relevant cue varying first?

Based on some preliminary work there are indications that the MAX-MIN order is better than MIN-MAX. The subject can account for more of the variation of the solution during the MAX. condition, and, it seems easier to build a complex concept when most of the variation is explained by the main construct.

METHOD

Task

The task used is the same as the one used by McHale and Stolurow (1962). Since the materials used and method of presentation are the same, they are not described in detail here. The stimuli consist of a red cross and a green cross presented in 2.5 inch by 2.5 inch squares. Each cross can appear in one of four horizontal and four vertical positions. The two relevant cues are the

horizontal positions of the crosses and the irrelevant cues are their vertical positions. The concept 'k' is defined as follows:

$$k = \frac{2x' + x''}{3}$$

where: x' is the horizontal position of the red cross
x'' is the horizontal position of the green cross
The positions are valued 3, 6, 9, and 12.

In this case the position of the red cross is weighted twice as much as that of the green cross.

The cue values for each stimulus of the asynchronous blocks are shown in Appendix A. The values of the fixed cues during each condition as well as the trial to trial values of the varying cues were chosen according to a table of random numbers. Since there were 16 trials and only 15 positions available for each condition, one stimulus appeared on two separate trials (for the MAX. condition, [6, 3] appeared on trials three and eight; for the MIN. condition [6, 12] appeared on trials two and eight).

The synchronous blocks are identical to those used by McHale and Stolurow (1962). The intercorrelations of the cues and 'k' are shown in Appendix B. On the synchronous blocks the stimuli were chosen so that the correlations of the irrelevant cues (y', y'') with 'k' were as close to zero as possible. Also, every combination of x' and x'' appeared once in each block of 16 trials so that the correlation of these cues was zero.

Procedure

Each group was given 160 presentations in five blocks of 32 trials. The first two blocks were the training trials and the last three were the task trials.

The 32 asynchronous trials consisted of 16 A-MAX trials and 16 A-MIN trials. The procedure is outlined in Table 1.

Table 1

Experimental Procedure

Group	Training Conditions (1-32)	Training Conditions (33-64)	Task Cond. (65-160)
Exp. 1	A(MIN-MAX)	S	S
Exp. 2	A(MAX-MIN)	S	S
Exp. 3	S	A(MIN-MAX)	S
Exp. 4	S	A(MAX-MIN)	S
Control	S	S	S

Each subject read the instructions shown in Appendix C. The following additional instruction was given verbally to each experimental group prior to the two types of asynchronous trials: "On the following presentations the red (green) cross will always appear in the same position. Therefore, changes in the value of 'k' will be caused by changed in the position of the green (red) cross."

After the second and the final blocks the subjects were given four test stimuli and asked to explain how they arrived at their answers.

Subjects

Subjects were taken from the Introductory Psychology course at the University of Illinois. There were 11 female subjects in each experimental

treatment. The control group consisted of eight females and three males chosen from a group which used identical materials except for the asynchronous training.

RESULTS

The Pearson product-moment correlations ("criterialities") of each subject's scores with each of the criteria (k, x', x", y', y") were calculated for the synchronous blocks and the two parts of the asynchronous block. For the purposes of analysis these were converted to Z' scores which are distributed in approximately normal form: (See Edwards,1960)

The average criterialities for each group are given in Appendix D. These are used in the graphs which follow, and were obtained by averaging the Z' scores and reconverting to a correlation.

The correlation of each subject's score with the correct score (k-criteriality) is used as the performance measure. The average k-criterialities are plotted in Fig. 1. The values for the asynchronous block are the averages of the values for the A-MAX. and A-MIN. portions. Part A of the figure shows the curves for the two A-S groups (Exp. I and Exp. II) and the control group while Part B shows the curves for the S-A groups (Exp. III and Exp. IV) and the control group.

<u>Analysis of Task Trials</u>

The experimental groups were compared using a 2 x 2 x 3 factorial design. The first factor involves the S-A vs A-S order comparison. The

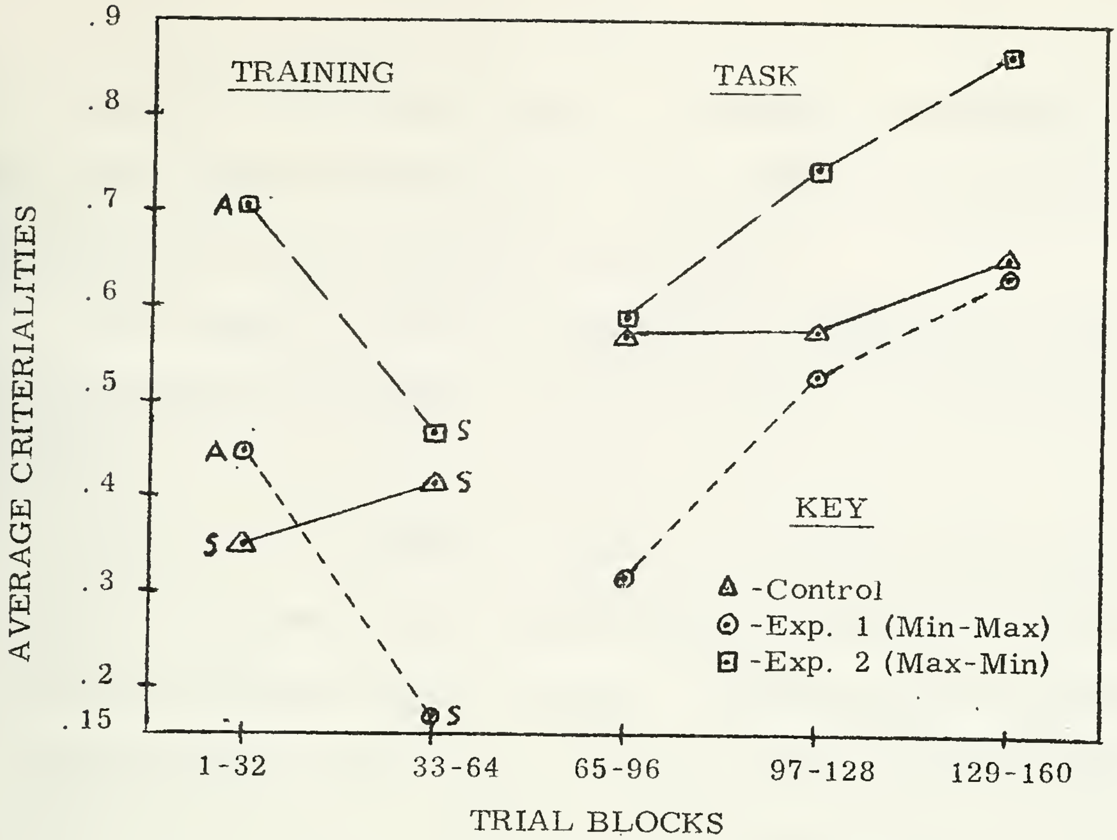


Figure 1A: Average Criterialities of the (A-S) Groups and the Control Group

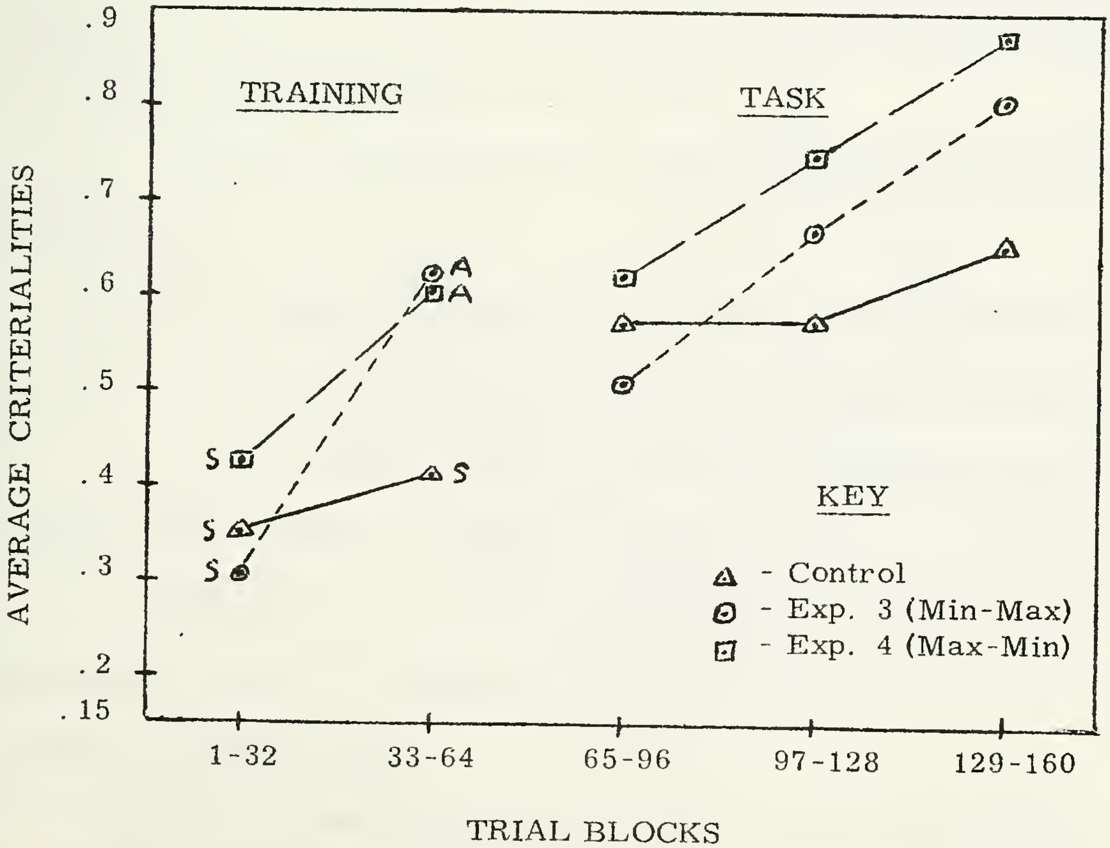


Fig. 1B. Average Criterialities of The (S-A) Groups and the Control Group

third involves the blocks factor which has the task blocks as its three levels. The analysis followed the procedure of Winer (1962) for a three-factor experiment with repeated measures on one factor.

The results are summarized in Table 2. No interactions were significant while the main effects of MAX-MIN order and trial blocks were significant.

The plan for comparing the control group with the experimental groups was as follows: first, the five groups were compared in a 5 x 3 repeated measures design on the three task blocks. If the overall group difference were significant,the control group would have been compared individually with each of the experimental groups. However, this was not the case as seen in Table 3. The group difference was not significant and the individual comparisons could not be made due to logical considerations

This analysis indicates the following:

1. There was an improvement in performance over the task trials for all groups.

2. There were no differences in the rate of learning among the groups.

3. The asynchronous training did not significantly improve performance.

4. It did not matter when the asynchronous training was given (i.e., initially or after a block of synchronous trials).

5. The groups who paid attention to the maximally pertinent cue first performed better than those who did not.

Analysis of Training Trials

Two comparisons were made among the experimental groups: first, performance on the asynchronous block was compared with performance on

Table 2

Summary of Analysis of Variance of Task Scores for the Experimental Groups[a]

Source	SS	d. f	MS	F
Between Subjects	33.383	43		
A(S-A Order)	.858	1	.858	1.188
B(MAX-MIN Order)	3.180	1	3.180	4.404[b]
A x B	.447	1	.447	.619
Subj. within groups	28.898	40	.722	
Within Subjects	18.112	88		
C(Trial Blocks)	6.781	2	3.391	24.396[c]
A x C	.050	2	.025	.180
B x C	.162	2	.081	.583
A x B X C	.021	2	.021	.079
C x Subj. within a groups	11.098	80	.139	

[a]The k-criterialities were transformed to Fisher-Z scores for this analysis.

[b]$F_{.95}(1, 40) = 4.08$

[c]$F_{.99}(2, 80) = 4.92$

Table 3

Summary of Analysis of Variance
of Task Scores for the
Experimental and Control Groups

Source	SS	d.f.	MS	F
Between Subjects	40.819	54		
A (Groups)	5.068	4	1.267	1.772
Subj. Within Groups	35.751	50	.715	
Within Subjects	19.309	110		
B (Trial Blocks)	6.051	2	3.026	22.917[a]
A x B	1.073	8	.134	1.015
B x Subj. within grps.	13.185	100	.132	

[a] $F_{.99}$ (2, 100) = 4.86.

the synchronous block; second, within the asynchronous block, the A-MIN and A-MAX trials were compared (see Fig. 2).

Each comparison was accomplished in a 2 x 2 x 2 factorial design with repeated measures on the third factor. The first two factors represent the arrangement of the experimental groups while the third factor represents the synchronous vs. asynchronous comparison in the first case and the A-MAX vs A-MIN comparison in the second case.

The summary for the S vs A comparison is given in Table 4. No interactions were significant while the main effects of MAX-MIN order and S vs. A comparison were significant.

The summary table for the A-MAX vs. A-MIN comparison is given in Table 5. The triple interaction was significant. Therefore, the comparison was made separately for each experimental group. The summary of this analysis is shown in Table 6. The performance difference on the two types of asynchronous trials was significant for Exp. -1 and Exp. -2 but not Exp. -3 and Exp. -4.

A partial check on the equivalence of the groups prior to training was made by comparing Exp. -3, Exp. -4 and the control group on Block 1 performance (see Fig. 1.). Since each of these groups received identical treatment prior to as well as on this block, any difference in their performances here would indicate that extra-experimental influences were affecting later observed differences.

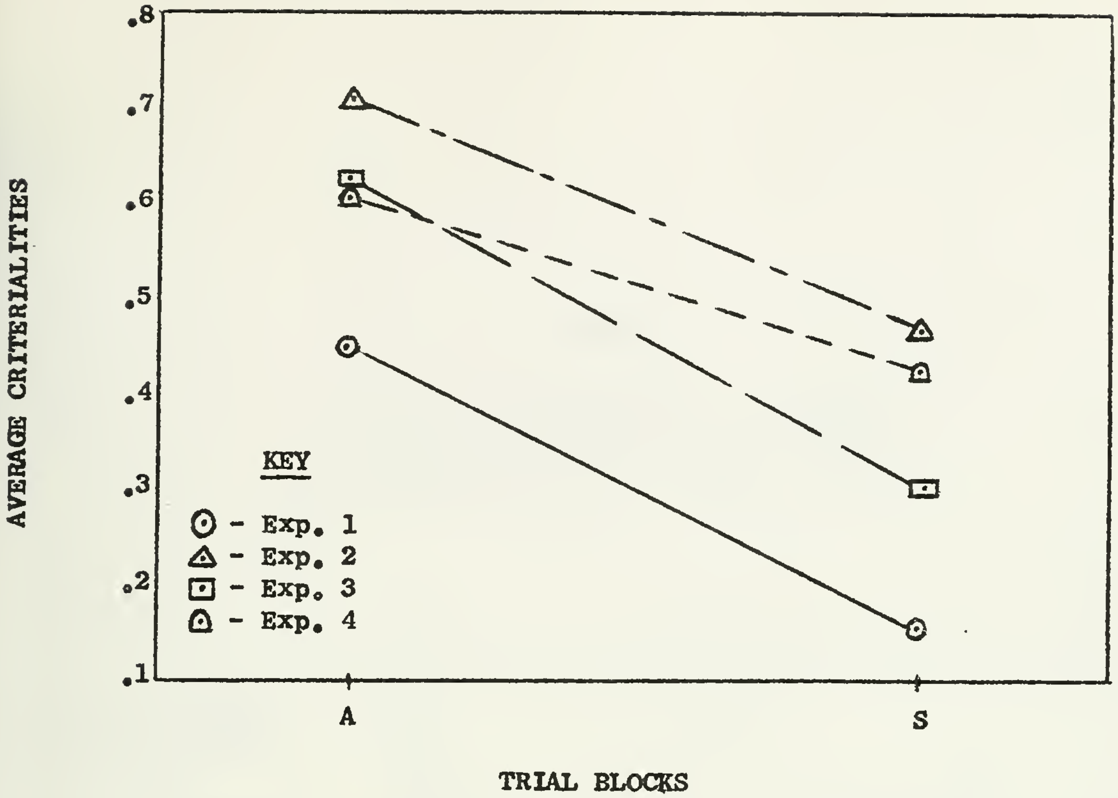


Fig. 2A. Synchronous vs. Asynchronous Performance on the Training Trails

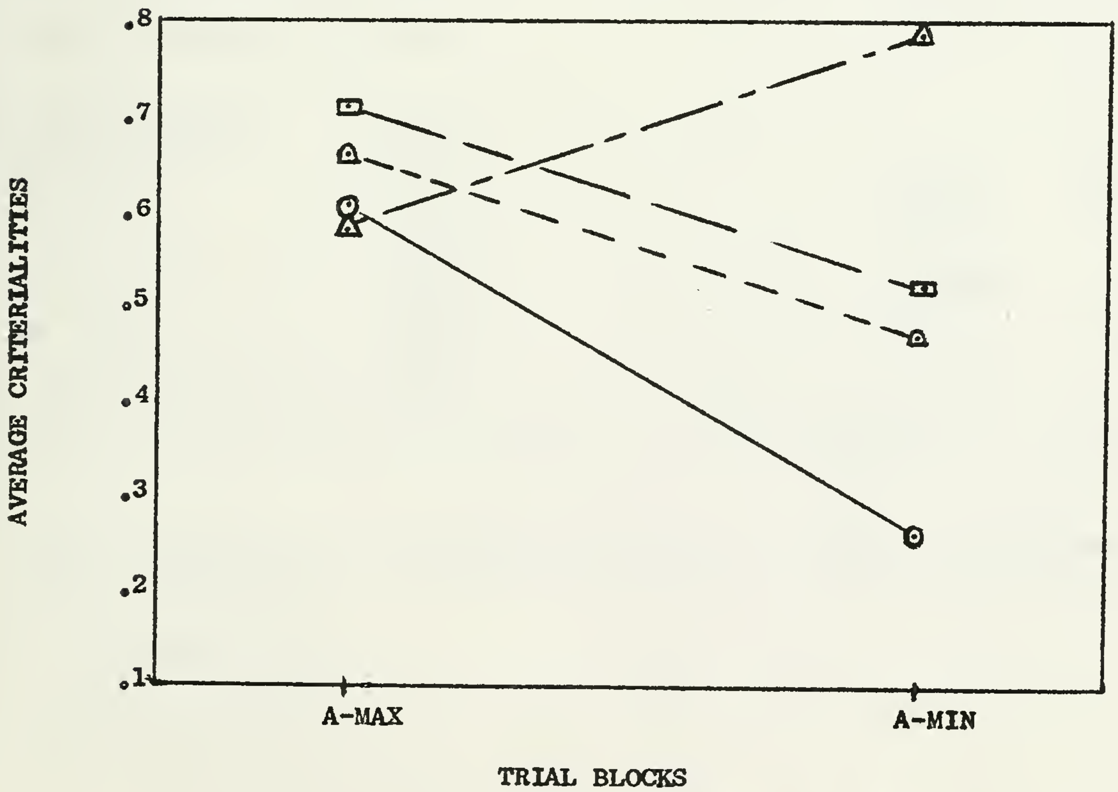


Fig. 2B. A-MAX vs. A-MIN Performance on the Training Tasks

Table 4

Summary of Analysis of Variance
For Synchronous vs. Asynchronous
Performance on the Training Trials[a]

Source	SS	d. f.	MS	F
Between Subjects	8.045	43		
A(S-A Order)	.025	1	.025	.154
B(MAX-MIN Order)	.893	1	.893	4.895[b]
A x B	.639	1	.639	3.944
Subj. within groups	6.488	40	.162	
Within Subjects	5.568	44		
C(S vs. A)	2.355	1	2.355	30.584[c]
A x C	.009	1	.009	.117
B x C	.045	1	.045	.584
A x B x C	.095	1	.095	1.234
C x Subj. within grps	3.064	40	.077	

[a]The Z' score for the asynchronous trials is the average Z' score for the A-MAX and A-MIN blocks.

[b]$F_{.95}(1, 40) = 4.08$

[c]$F_{.99}(1, 40) = 7.31$

Table 5

Summary of Analysis of Variance
For A-Max. vs. A-Min. Performance
on the Training Trials

Source	SS	d. f.	MS	F
Between Subjects	13.217	43		
A(S-A Order)	.004	1	.004	.013
B(MAX-MIN Order)	.537	1	.537	1.878
A x B	1.229	1	1.229	4.297
Subj. within groups	11.447	40	.286	
Within Subjects	10.115	44		
C(A-MIN vs. A-MAX)	.516	1	.516	2.851
A x C	.434	1	.434	1.845
B x C	1.026	1	1.026	5.669[a]
A x B x C	.905	1	.905	4.448[a]
C x Subj. within groups	7.234	40	.181	

[a] $F_{.95}\ (1, 40) = 4.08$

Table 6

Summary of Analysis of Simple Effects for A-Min. vs. A-Max. Performance on the Training Trials[a]

Source	SS	d. f.	MS	F
Within Subjects	10.115	44		
C at A_1B_1(MIN-MAX, S)	1.024	1	1.024	5.657[b]
C at A_1B_2(MAX-MIN, S)	.906	1	.906	5.006[b]
C at A_2B_1(S, MIN-MAX)	.518	1	.518	2.860
C at A_2B_2(S, MAX-MIN)	.432	1	.432	2.386
C x Subj. within groups	7.234	40	.181	

[a] See Table 5 for basic analysis relating to triple order interaction and explanatio of symbols. See Fig. 2B for graph.

[b] $F_{.95}(1, 40) = 4.08$

The method of analysis follows the procedure of Winer (1962, p. 65) and is summarized in Table 7. Although there was an overall significant difference for the five groups, there were no significant differences among the three groups which received identical treatment.

The analysis of the training trials indicates that:

1. Performance on the A-block was highly superior to performance on the S-block regardless of the order in which they were given.

2. The MAX-MIN groups performed better than the MIN-MAX groups.

3. The A-MAX vs. A-MIN comparison was different for the various experimental groups. Exp. -1 performed significantly better on the A-MAX trials while Exp. -2 performed significantly better on the A-MIN trials. Both Exp. -3 and Exp. -4 performed better on the A-MAX trials but the difference was not significant. (Fig. 2B).

4. Exp. -3, Exp. -4 and the control group can be considered identical in terms of initial performance on this task.

Cue Criterialities

The average cue criterialities for each of the groups are plotted in Figures 3 to 7. Also plotted in each figure are the k-criterialities. The value for each cue on the asynchronous blocks was taken from the portion (A-MAX or A-MIN) on which that cue was free to vary.

Table 7

Summary of Analysis of Variance
For the Experimental Groups
and the Control Group on the First Trial Block

Source	SS	d. f.	MS	F
Treatment	2.151	4	.538	5.38[a]
Exp. 3 vs Exp. 4	.111	1	.111	1.11[b]
Exp. 3 vs Control	.014	1	.014	.14
Exp. 4 vs Control	.046	1	.046	.46
Error	5.000	50	.100	
Total	7.150	54		

[a] $F_{.99}(4, 50) = 3.74$

[b] $F_{.75}(1, 50) = 1.35$

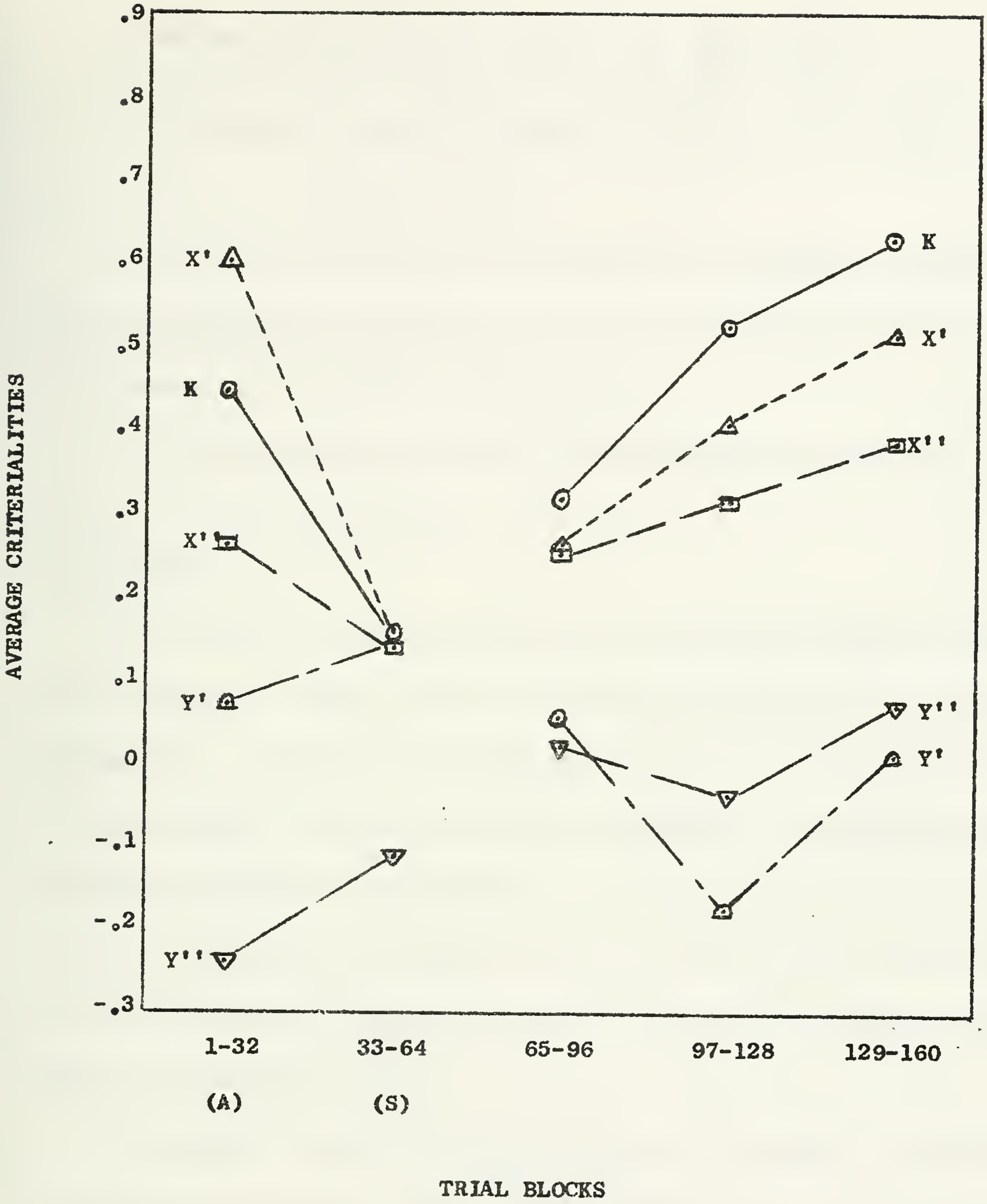


Fig. 3. Criterialities for Experimental Group 1 (MIN-MAX S)

In addition to the previous findings these indicate that:

1. All groups learned to ignore the irrelevant cues y' and y".
2. There was a general decrease in performance in going from an asynchronous to a synchronous block.
3. The gain in performance obtained when going from a synchronous block to an asynchronous block was greater than that obtained when going from a synchronous to another synchronous block.
4. The experimental groups weighted the relevant cues (x', x") more heavily than did the control group.

Questionnaire

After the 64th trial and again at the end of the experiment, the subjects were asked to assign 'k' values to four test stimuli and to indicate how they arrived at their answers. The amount of information which was obtained in this way was lexx than expected. The subjects found it difficult to verbalize their strategies or methods of solution. In fact, there were cases where S was close to the solution in terms of the performance measure (k-criteriality >.85) and yet was able to express only a vague notion about the fact that the crosses appeared in columns with different weights.

Early hypotheses were generally specific to some area of the square; e.g., : "if both crosses are at the left side, k=3; if both crosses are at the right side, k=12." The notion of columns seemed to enter first, followed by the idea of differing weights for the two crosses. Also, subjects who had relatively sophisticated strategies, generally used the red cross as the main factor;e.g., : "get distance of the red cross from the right or left side of the square and then the distance of the green cross from the red cross."

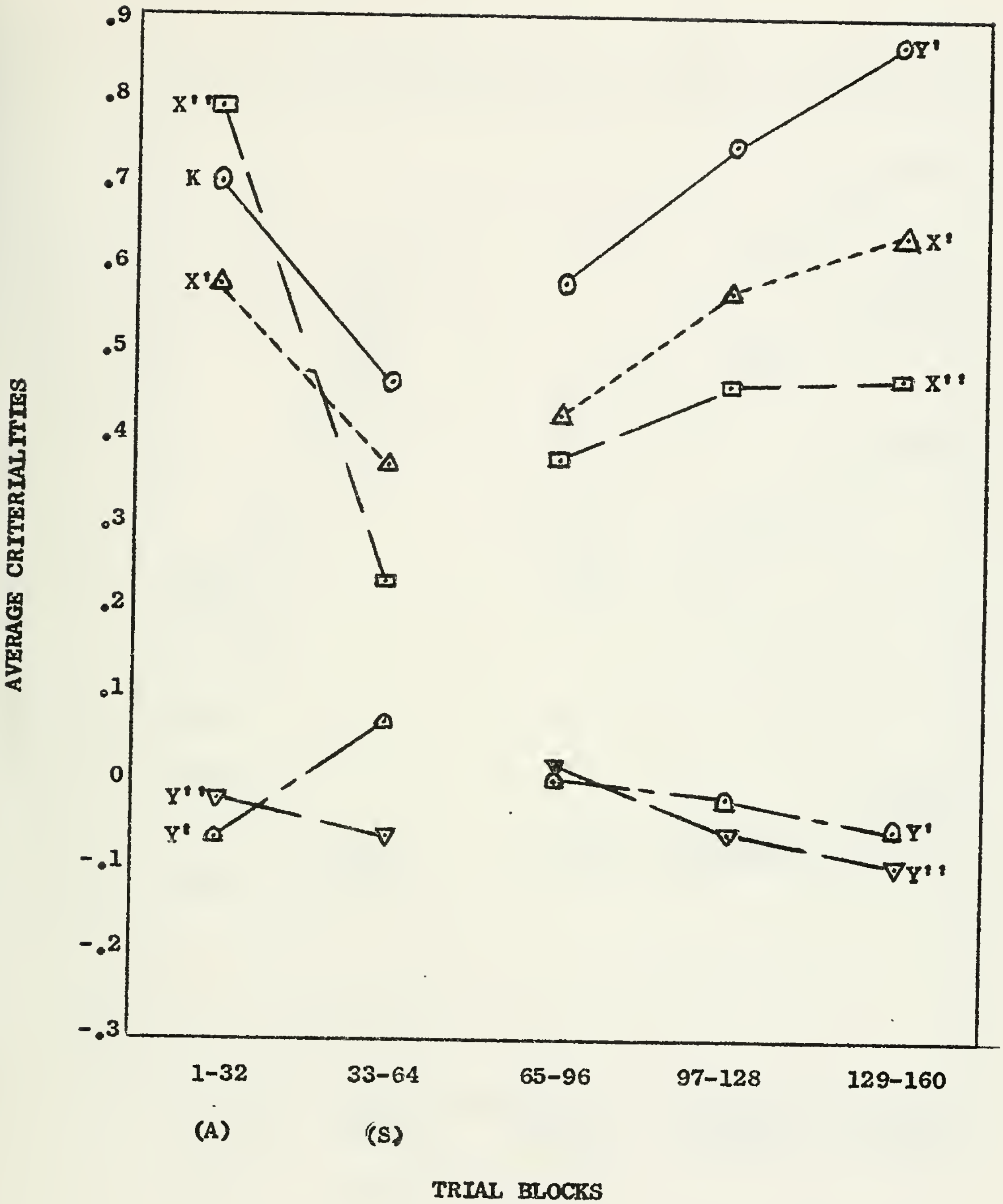


Fig. 4. Cue Criterialities for Experimental Group 2 (MAX-MIN, S)

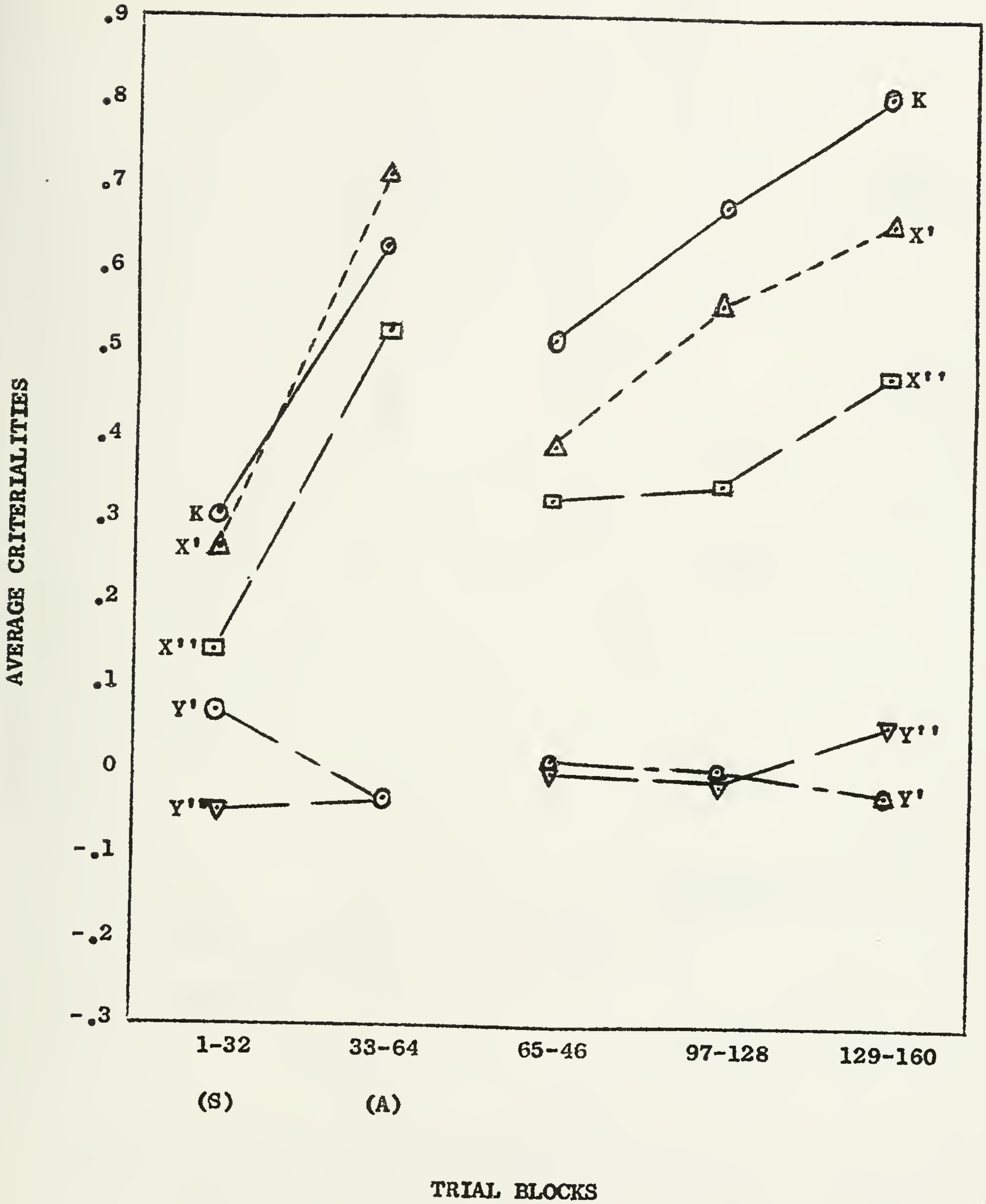


Fig. 5. Cue Criterialities for Experimental Group 3 (S, MAX-MIN)

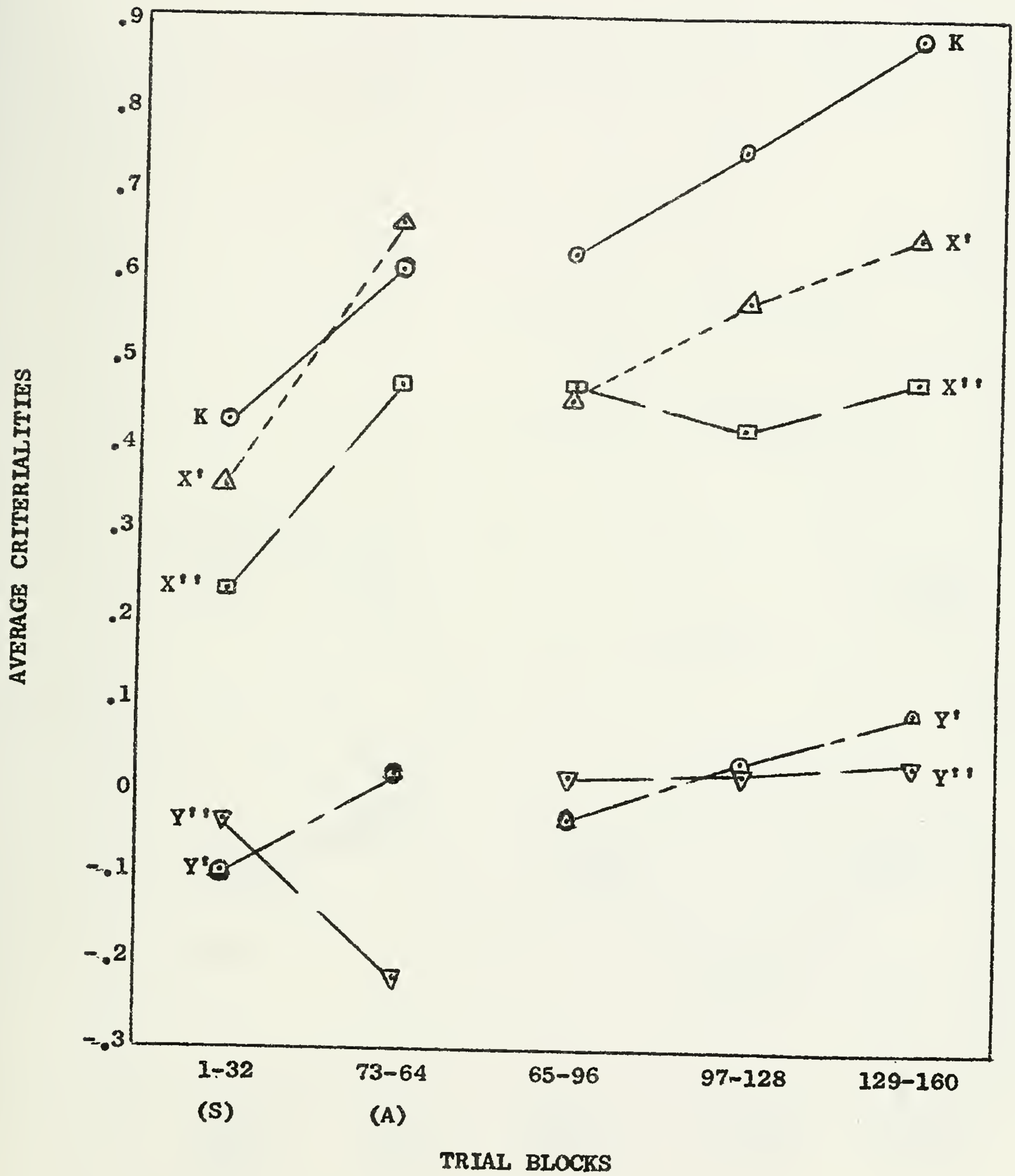


Fig. 6. Cue Criterialities for Experimental Group 4 (S, MAX-MIN)

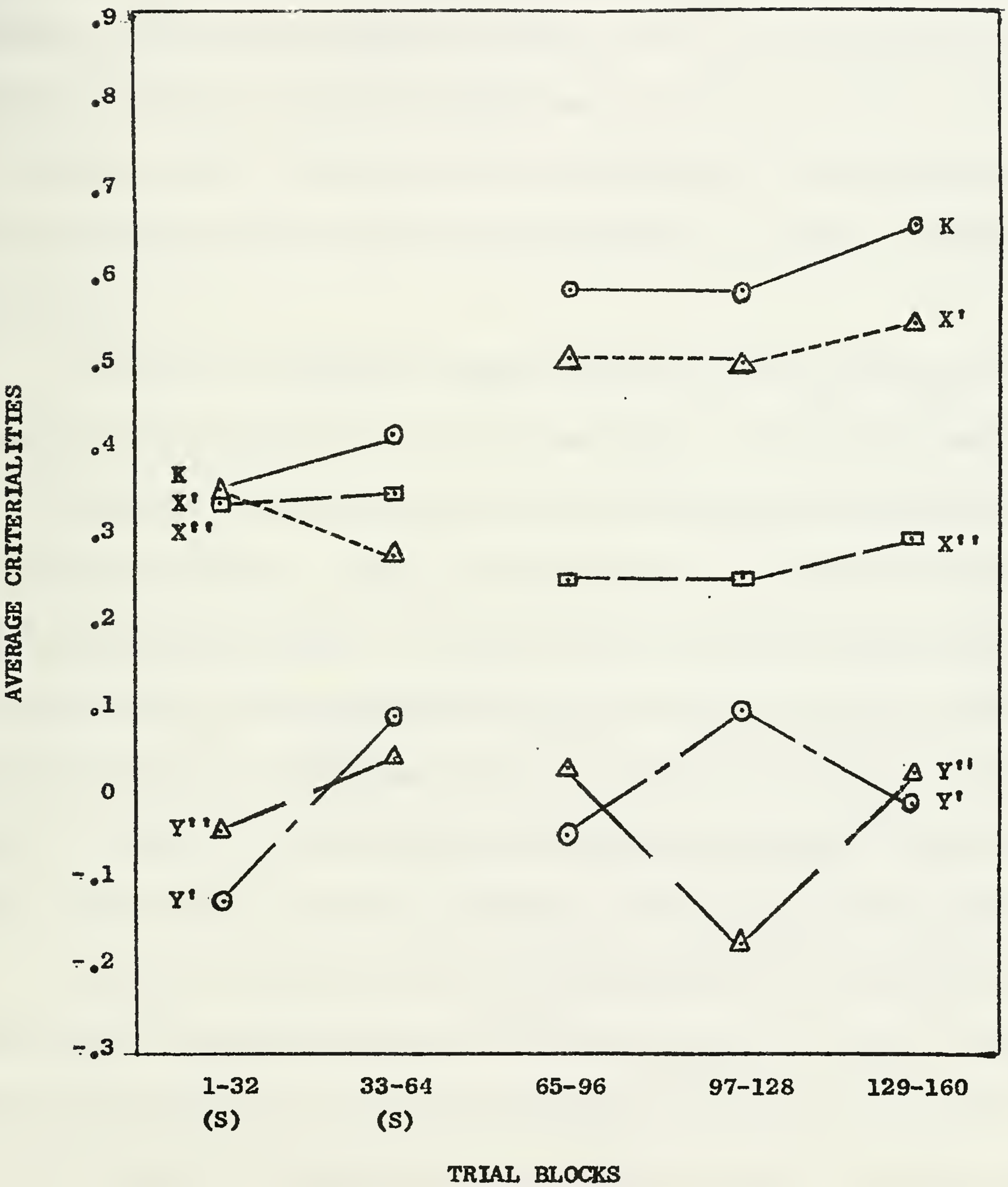


Fig. 7. Control Group Criterialities

DISCUSSION

The discussion is organized according to the questions listed on pages 2 & 3.

1. What is the effect of adding the asynchronous trials?

As expected, the performance on asynchronous trials was far superior to performance on synchronous trials during training (Table 4). This agrees with the findings of Detambel and Stolurow. However, there is no clear indication that the asynchronous groups performed better than the synchronous group on the transfer or task trials. Fig. 1 shows three of the experimental groups performing above the control group on the last block, but this difference was not statistically significant.

The lack of a significant improvement in task trial performance as a result of asynchronous training could be due to one of the following factors. First, the control group consisted of eight females and three males while the experimental groups had females only. Also, the control group experiment was administered by a different experimenter. It is felt, however, that these differences in procedure did not have a significant effect on the results. This is partially verified by the lack of differences among the three S groups on the first block (Table 7). However, the possibility exists that males perform better on this task, giving the control group an advantage over the experimental groups.

Second, if the subjects use relationships between the crosses, then allowing them to concentrate on one at a time should be of no help. From the questionnaire results it seems evident that the subjects generally used the position of each cross separately to arrive at their answers. Therefore, this explanation is tentatively discarded.

Third, this study can be discussed in terms of the single-problem training versus multiple-problem training issue. Morrisett and Hovland (1959) have shown that there are two factors which seem necessary to produce a high degree of transfer in multiple-problem training: (1) a high degree of learning must be achieved on each single problem; (2) several problems must be presented to insure learned generalization. Although the first factor was satisfied as indicated by asynchronous performance, it seems that the second was not. Each of the two types of asynchronous trials used could be considered to have been taken from a set of 16 types of asynchronous blocks (one for each of the 16 possible positions of the fixed cross). According to the findings of Morrisett and Hovland, training on several members of these sets is necessary before significant transfer can be obtained. Since the groups in this study were trained on only one member of the set of 16 possible A-MAX blocks and one member of the set of 16 possible A-MIN blocks, then transfer to the total problem situation was low.

Fourth, there is some evidence that the experimental groups tended to overweight the green cross or less pertinent cue in arriving at their results (see Fig. 3-7). All of the experimental groups had much higher x''-criterialities than the control group on the last block. Also, for three of the experimental groups, the x''-criteriality was greater than the validity of x'' (the correlation of x'' with k, which is about .45 on each block). This is evidence (although not conclusive in a statistical sense) that the A-MIN condition was actually detrimental in that it led to the over-use of the less pertinent cue.

The conclusion is that the asynchronous training did not significantly improve performance on the transfer task as a result of two factors:

(1) the lack of generalized learning set due to training on only one member of the class of asynchronous blocks.

(2) the over-emphasis of the less pertinent cue as a result of asynchronous training trials on which that cue was allowed to vary.

This is only a tentative conclusion since it results from only one study. However, it will be used as a working hypothesis for further research in this area.

2. What is the best order of presentation of the asynchronous (A) and synchronous (S) training trials?

As shown in Table 2 there is no significant difference due to the two orders of presenting asynchronous and synchronous blocks during training. It was expected that the S-A order would have an orienting effect similar to the effect of introductions and overviews prior to presenting detailed verbal reports. However, this was not demonstrated in this experiment.

The conclusion, then is that synchronous training prior to asynchronous training does not improve efficiency of transfer. Again, this is tentative, since the amount of prior synchronous training is probably critical. This factor should be examined carefully before an overall conclusion is made concerning the effect of initial synchronous training.

3. During the asynchronous trials, is it better to present the more relevant cue varying first?

The answer to this question is "yes" since the groups which received A-MAX training first performed significantly better on the task trials than those which received A-MIN training first. This indicates that the order of training on complex tasks should proceed from the most pertinent aspects to the least pertinent aspects. Also, the second conclusion in the discussion of question 1 gives rise to the possibility that the amount of training should be lower for the less pertinent aspects.

FURTHER RESEARCH

The conslusions arrived at in this experiment will serve as the basis for future investigations. In particular, the following questions are of interest:

1. How does the <u>type</u> of asynchronous training affect transfer:

On the basis of this experiment it is expected that training which involves only maximally pertinent cues should be superior to that which involves only minimally pertinent cues.

2. How does the <u>amount</u> of asynchronous training affect transfer?

There are several issues here. First, there is the question of how much training should be given for each type of asynchronous block (fixed cross in a given position). Second, there is the question of how many types of asynchronous blocks should be presented for optimum transfer. These two questions relate to the multiple problem training issue as discussed by Morissett and Hovland.

Third, there is the question of apportioning training among the more pertinent aspects: Is it better to decrease the amount of training for the less pertinent aspects of a problem as compared with the more pertinent aspects? There is some slight evidence that this is the case. However, a direct test of this hypothesis is necessary before a definite conclusion can be made.

3. How does the <u>order</u> of asynchronous training affect transfer?

It was demonstrated in this experiment that transfer is greater when the more pertinent cue is allowed to vary first. This suggests that there is an order relationship in training which is based on the relevancy of the aspects. That is, the more relevant, or pertinent aspects should be presented first.

This should be demonstrated for the case of three or more aspects, each differing in their relevancy to the solution before this order effect is accepted.

SUMMARY

Several ways of structuring and sequencing the early trials of a complex task were compared. Four experimental groups received both structured (asynchronous) and unstructured (synchronous) training trials. The asynchronous trials were divided into two segments: A-MAX (the more relevant cue was free to vary) and A-MIN (the less relevant cue was free to vary.) The four experimental conditions were generated by the different sequential orders of presenting the structured and unstructured trials (A-S vs. S-A) as well as the two types of asynchronous trials (MAX-MIN vs. MIN-MAX). A control group received only synchronous training.

It was found that asynchronous training did not significantly improve performance. It is felt that this lack of improvement was due to the following two factors:

(1) the lack of generalized learning due to training on only one member of the class of asynchronous blocks.

(2) the over-emphasis of the less relevant cue.

For the experimental groups it was found that the presentation of synchronous training trials prior to asynchronous training did not improve performance. Therefore, the hypothesis that this would aid transfer by familiarizing subjects with the transfer task prior to synchronous training was not founded.

It was found that presenting a sequence in which the maximally pertinent cue varied first (A-MAX condition) led to improved transfer task performance. This is an indication that the order of training on a complex task should proceed from the more relevant to the less relevant aspects.

REFERENCES

Azuma, H. Comparison of a correlational with a probabilistic approach to concept learning. Urbana, Ill.: Univer. of Ill., Training Res. Lab., 1960. Unpublished doctoral dissertation.

Detambel, M. H. and Stolurow, L. M. Stimulus sequence and concept learning. J. exp. Psychol., 1956, 51, pp. 34-40.

Edwards, A. L. Experimental design in psychological research. New York: Rinehart & Company, Inc., 1960.

McHale, T. J. and Stolurow, L. M. Concept formation with metrically multi-valued cues and response categories. Urbana, Ill: Univer. of Ill., Training Res. Lab., Tech. Rep. No. 3, ONR 1834(36), April, 1962.

Morrisett, M. L. and Hovland, C. I.. A comparison of three varieties of training in human problem solving. J. exp. Psychol., 1959, 58, pp. 52-55.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Appendix A

Cue Values During the Asynchronous Blocks

	Max Condition				Min Condition			
Trial	x'	y'	x''	y''	x'	y'	x''	y''
1	3	12	3	6	6	9	12	12
2	9	6					6	12
3	6	6					3	6
4	6	3					12	9
5	12	3					6	3
6	9	12					6	6
7	3	3					12	6
8	6	6					6	12
9	12	12					9	3
10	6	12					9	12
11	9	3					3	3
12	12	9					3	9
13	6	9					9	6
14	12	6					9	9
15	3	6					12	3
16	9	9					3	12

Appendix B

Intercorrelations of the Cues and K

A. Training Trials

Block

A-MAX

	x'	y'	x''	y''
k	1	-.026	0	0
x'		-.026	0	0
y'			0	0
x''				0

A-MIN

	x'	y'	x''	y''
k	0	0	1	-.024
x'		0	0	0
y'			0	0
x''				-.024

S_1

(first block of the control group)

	x'	y'	x''	y''
k	.894	-.050	.447	.117
x'		-.023	.000	.113
y'			-.138	.047
x''				.038

S_2

(S block for the Exp. Groups, and S_2 block for the control)

	x'	y'	x''	y''
k	.894	.039	.447	-.034
x'		.013	.000	-.025
y'			.063	-.087
x''				-.025

B. Task Trials

S_3
(65-96)

	x'	y'	x"	y"
k	.894	-.057	.447	.112
x'		-.051	.000	.100
y'			-.-26	-.077
x"				.050

S_4
(97-128)

	x'	y'	x"	y"
k	.894	.066	.447	.022
x'		.046	.000	-.025
y'			.055	.114
x"				.100

S_5
(129-164)

	x'	y'	x"	y"
k	.894	.045	.447	.047
x'		.000	.000	.000
y'			.100	.026
x"				.105

Appendix C

INSTRUCTIONS TO STUDENTS

This task is designed to study your ability to learn to use relevant information to evaluate something. We will obtain a score showing how rapidly and how well you are able to use information in solving a problem which you do not completely understand from the beginning. It is to make evaluations we will call k-ness.

We are going to show you some drawings, and we want you to judge how much k-ness each figure has. We will tell you how to compute k later.

Turn to page one which has six figures on it. Each figure has a red cross and a green cross. The position of the red cross and green cross determine the value of k for each figure. Look at your answer sheet. For each item, there are ten possible answers: 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12. You are to figure out how to predict the k-scale value from information given you in each figure.

We will show you a series of figures, like the ones you see on page one. You are to decide the value of each on the scale; that is, you are to decide whether its k-value is 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12.

<u>HOW TO USE YOUR ANSWER SHEET</u>:

For each drawing there is one row on your answer sheet. That row contains ten numbers, which are the ten possible responses: 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12. After you decide how large k is for a figure, mark your answer by drawing an X through the proper number. You will be allowed 20 seconds for each figure. Twenty seconds sounds like too short a time, but you will find that it is sufficient. At the end of each 20-second period, I will tell you the correct answer for that item. You may underline the correct answers if you wish. Do not draw another X through the correct answer.

<u>Remember:</u>

1. You are to judge the k-value of each figure.

2. You then draw an X through the proper number on your answer sheet. You will be told the correct answer at the end of each 20-second interval. Ideally, you should be able to answer each item correctly.

Do not look at any page in the booklet of figures except the one you are working on at the moment. That is, do not look back or ahead. If you finish an item before the allotted 20-second interval is over, wait until its answer is given before proceeding on the next item.

Wait for the signal to begin.

Appendix D

Average Criterialities

Exp. 1 (MIN-MAX, S)

BLOCKS

		1-16 A-MIN	17-32 A-MAX	33-64 S	65-96 S	97-128 S	129-160 S
	k	.261	.603	.155	.318	.523	.632
	x'	0	.603	.152	.261	.408	.511
variable	y'	0	.071	.140	.053	-.172	.010
	x''	.261	0	.145	.253	.315	.388
	y''	-.436	0	-.110	.022	-.040	.069

Exp. 2 (MAX-MIN, S)

		1-16 A-MAX	17-32 A-MIN	33-64 S	65-96 S	97-128 S	129-160 S
	k	.586	.793	.468	.589	.743	.863
	x'	.586	0	.378	.432	.577	.643
variable	y'	-.061	0	.074	-.049	-.015	.005
	x''	0	.793	.239	.379	.464	.476
	y''	0	-.021	-.053	-.099	-.055	.016

Exp. 3 (S, MIN-MAX)

		1-32 S	33-48 A-MIN	49-64 A-MAX	65-96 S	97-128 S	129-160 S
	k	.304	.522	.710	.511	.672	.803
	x'	.268	0	.710	.390	.559	.658
variable	y'	.071	0	-.036	.015	.003	-.023
	x''	.148	.522	0	.327	.343	.476
	y''	-.046	-.031	0	.001	-.010	.062

Exp. 4 (S, MAX-MIN)

BLOCKS

	1-32 S	33-48 A-MAX	49-64 A-MIN	65-96 S	97-128 S	129-160 S
k	.427	.660	.472	.621	.747	.878
x'	.355	.660	0	.453	.568	.648
variable y'	-.087	.023	0	-.033	.036	.097
x''	.232	0	.472	.461	.421	.480
y''	-.032	0	-.218	.016	.028	.038

Control Group

	1-32 S	33-64 S	65-96 S	97-128 S	129-160 S
k	.350	.413	.578	.574	.652
x'	.351	.272	.500	.494	.540
variable y'	-.124	.089	-.050	.096	-.009
x''	.336	.343	.244	.246	.297
y''	.046	.041	.229	-.174	.020

TRAINING RESEARCH LABORATORY

Department of Psychology   Bureau of Educational Research
University of Illinois   8 Lincoln Hall   Urbana, Illinois

MORE INFORMATION -- CUES OR PRINCIPLE?

Thomas J. McHale and Lawrence M. Stolurow

Technical Report No. 5

Psychological and Educational Factors
in Transfer of Training
Phase I

May, 1964

U. S. Office of Education
Contract 2-20-003

Lawrence M. Stolurow
Principal Investigator

TRAINING RESEARCH LABORATORY
Bureau of Educational Research
Department of Psychology
University of Illinois
Urbana, Illinois

MORE INFORMATION -- CUES OR PRINCIPLE?

Thomas J. McHale and Lawrence M. Stolurow

Technical Report No. 5

May, 1964

Psychological and Educational Factors in Transfer of Training
Phase I

Principal Investigator:
Lawrence M. Stolurow

Project Sponsor:
Educational Media Branch
U. S. Office of Education
Title VII
Contract No. 2-20-003

Table of Contents

List of Tables

List of Figures

MORE INFORMATION -- CUES OR PRINCIPLE?

Thomas J. McHale[1] and Lawrence M. Stolurow
University of Illinois

Concern for the importance of stimulus factors in concept formation can be traced to Hull's classic study (1920). Later Smoke (1933) studied the relative importance of positive and negative instances, and Hovland (1952) re-examined this problem in terms of information theory. More recently, the work of Brunswik (1956) stimulated renewed interest in the problem. Bruner, Goodnow, & Austin (1956) report a series of studies concerned with the way in which S learns to select and utilize cues. These authors coined the term "criteriality" as a measure of the degree to which the S uses a particular cue in forming his responses; however, since they used two-valued cue and response categories (e.g., swept-back wing or delta wing, an X plane or a non-X plane), they were unable to infer from S's responses the nature of the mediating construct or principle being used. These should not be interpreted as the defining conditions for the use of the

---

[1]Now Assistant Professor of Psychology at Gonzaga University, Spokane, Washington.

term, however, since another investigator, Smedslund (1955), utilized scaled cue and scaled response categories so that E could give scaled feedback rather than mere confirmation or disconfirmation. He analyzed his data in terms of the criterialities of each of the various cues in order to determine the use that was being made of various hypotheses or principles. Since his task was quite complex, the likelihood of success for the S was limited and little information was gleaned from the procedure.

Azuma (1960), recognizing the intention of this approach and the shortcomings of these efforts to implement it, developed a task with metrically multi-valued cue and response categories based on a multiple-correlation model which, although complex, could be mastered more easily than the task used by Smedslund. Azuma's study compared two task models: (a) a multiple-correlation model and (b) a decision theory model involving probability learning conditions. Both tasks were specifically designed to reveal the nature of the mediating principle, used by S, through an analysis of the various cues' criterialities. There were four cues, two relevant and two irrelevant. With the proper linear combination of the two relevant cues, an S could always determine the value of k on every trial. S had to determine which cues were relevant and the proper linear combination to weight them to arrive at the value of k, the unknown. The data from this study were submitted to further analyses (Azuma & Cronbach, 1961, 1962; Cronbach & Azuma, 1961).

McHale & Stolurow (1962), in a replication of one part of the Azuma study, attempted to determine how information about the number of relevant cues improves performance. The information was not effectively presented and the number of subjects was small; the resulting trend, while in the right direction, did not reach significance. This suggested the need for a more complete investigation of the amounts of information carried by various parts of the mediating construct. Thus,they shifted the problem to the question of the amount of information conveyed to S by different parts of the mediating construct presumably used by S in arriving at the correct k values.

Problem

The present study was designed to investigate the amounts of information communicated by two components of a quantitative task: (a) knowledge of a principle, and (b) knowledge of the number of critical cues. It also investigated the relative effectiveness of knowledge of principle as opposed to knowledge of cues at different stages of learning. Since there was some question as to whether or not complete knowledge of the task solution would lead to perfect performance, a full information group was included; in addition,a no information group was included to determine the lower boundary of performance.

Design

The study was a 2 x 2 design generated by two dichotomized variables (presence or absence of information about cues, and presence or absence of information about principles).

The four groups consisted of (a) a cue group, who knew the set of four possible cues and the number of required cues; (b) a principle group, who knew only the principle; (c) a full information group, who knew both cues and principle; and (4) a no information group, who knew neither cues nor principle.

Hypotheses

An analysis of the task relative to the possible hypotheses which S could consider led to the following specific hypotheses which were tested:

1. Knowledge of the principle would be more beneficial than knowledge of the number of critical cues since there appear to be many more possible ways of weighting or combining cues than there are possible cues.

2. Knowledge of the number of critical cues would be more beneficial in the early stages of learning when the S must detect what is relevant, and knowledge of the principle would be more beneficial in the later stages of learning when the S must determine the appropriate weights to use in combining the relevant cues to determine k.

3. The rank order of performance of the four groups would be as follows: full information, principle information,

cue information, and no information.

Method

Subjects

Fifty-two undergraduate students in psychology at the University of Illinois participated in the experiment with 13 subjects in each of four groups. Thirty-eight subjects were administered the task during a regular class period; the other 14 subjects were obtained from a subject pool and were administered the task in small groups. Of the latter, 4 were in the cue group, 4 in the full information group, and 6 in the principle group.[2]

Materials and Procedure

A booklet of stimulus presentations, answer sheets, and a questionnaire were distributed to all Ss before instructions were given. Depending upon the group to which the S belonged, one of four sets of instructions was then read by each S.

The task stimuli. Each stimulus (trial) consisted of a 2.5 inch by 2.5 inch square outline with a small red cross

[2]The classroom administration obviously was not an optimal situation since the data for eleven subjects from the undergraduate class who refused to cooperate had to be discarded.

and a small green cross drawn inside it. The left side and the bottom of the square represented coordinate axes. The location of each cross was specified by its coordinate distances from the left side and the bottom of the square,

there were four possible positions along each coordinate.

Each of the crosses can be represented as a letter, x or y, and the four coordinate values can be represented by four variables, x', y', x", and y". Each variable could take on one of four values, .5, 1.0, 1.5, or 2.0 inches. The number of possible combinations of these values was 4^4, or 256; however, since the crosses were not allowed to occupy the same location in any stimulus presentation, only 240 (16 x 15) combinations were actually possible and not all of the possible displays were used.

Presentation of stimuli. Stimulus displays were presented in a booklet in which each page contained six different displays. The booklet consisted of 128 stimuli, or trials, to which Ss responded by marking one of ten possible response categories with an X. The answer sheet contained 10 circles for each trial, each circle was for one of the 10 possible numerical answers and S drew an X through the appropriate circle to indicate his response, i.e., what he thought was the value of k for the display. After each trial S was told the true value of k.

The 128 learning trials can be considered as 8 sets of 16 different presentations. Within each set, the 16 possible combinations of x' and x" appeared once each, which automatically made $r_{x'x''}$= .00. The distributions of y' and y" were very close to rectangular. Displays were prepared so that, for the set, $r_{x'y'}$, $r_{x'y''}$, $r_{x''y'}$, and $r_{x''y''}$ did not exceed .12 in any block. Thus, for practical purposes, these variables can be considered to be uncorrelated.

Criterion k. The formula used by the E to define the oorrect response, k, was (2x' + x")/3. Since x', y', x" and y" were uncorrelated with each other, the definition of k, in terms of the zero order correlations, determined their validities as follows: $r_{x'k}$=.89, $r_{x''k}$=.45, $r_{y'k}$=.00, $r_{y''k}$=.00. Though the actual correlations of x' and x" with k were exactly .89 and .45, respectively, in each block, the actual correlations of y' and y" with k varied between -.12 and +.12. Since the 10 discreet response categories were exact (except for rounding in the second decimal place), the Ss had to use precisely a 2:1 weighting in order to be correct 100% of the time.

Measures of Performance

The dependent variables used as measures of performance were the criterialities of the individual cues, computed over blocks of trials, and the value given the construct k .

Product-moment correlation coefficients were computed for each S's actual responses with the Ss possible responses, based upon the assumption that he made judgments solely in terms of x', x", y', y", or k. This yielded a 5 x 4 matrix of correlations for each subject (rows for x', x", y', y", and k; columns for each block of 32 trials) which was analyzed separately. Correlations to determine criterialities were computed over nonoverlapping blocks of 32 trials: 1-32; 33-64, 65-96, and 97-128. These blocks will be referred to as blocks 1, 2, 3, and 4 respectively.

A questionnaire which attempted to get S to verbalize his principle or mediating construct and to assign a number to his relative weighting of the two relevant cues was administered to all Ss in all groups after the learning session was completed. The verbalizations are informative, although one question about the relative weighting was not understood by Ss (See discussion).

Results

In order to determine representative values for each group, the following steps were taken:

1. Criterialities were converted into z' scores and group means for the z' scores were calculated. All statistical tests used the z' transformations as raw scores since they roughly approximated a normal distribution.

2. The mean *z'* values were reconverted into the corresponding *r* values. These mean *r* values were taken as the representative criterialities for each group in each block of trials (mean criterialities). Table 1 contains the mean criterialities on all coordinates for each experimental group on each of the four trial blocks. Figures 1, 2, and 3 are graphical representations of the data in Table 1 (graphical representations of the mean criterialities of y' and y" are not presented because the values are close to .00, as predicted). Note that, in general, an asymptote for the learning curves of *k* has not been reached.

A 2 x 2 analysis of variance was used for x', x", and *k* to detect the significant effects of principle, cues, or their interactions for each of the 4 trial blocks. In addition, Duncan's Multiple Range Test (1955) was used to test for significant differences between the means of the groups in each of the 4 blocks. Tables 2 through 6 contain the results of these tests.

Table 1

Mean Criterialities ($\underline{r}$) for Each Group in All Blocks of Trials

Group[a]	Block	Mean criterialities[b]				
		x'	y'	x"	y"	$\underline{k}$
No Information	1	.38	-.03	.22	.03	.42
	2	.64	.00	.31	.07	.71
	3	.64	-.12	.33	.08	.74
	4	.69	.04	.36	.06	.78
Cue Information	1	.40	-.03	.26	.15	.49
	2	.56	.07	.43	.05	.73
	3	.66	-.08	.52	.09	.86
	4	.66	.10	.53	.09	.88
Principle Information	1	.38	.00	.31	.03	.51
	2	.54	.04	.45	.01	.71
	3	.68	-.09	.35	.00	.83
	4	.66	-.01	.36	-.04	.79
Full Information (cue and principle)	1	.69	-.04	.44	.08	.85
	2	.78	.11	.35	-.03	.92
	3	.77	-.04	.40	.11	.94
	4	.79	.04	.38	.00	.95

[a]N was 13 per group; therefore any r which was greater than .55 exceeded the .05 level of significance (two-tailed test).

[b]Criterion criterialities are $r_{x'\underline{k}}=.89$; $r_{x''\underline{k}}=.45$, $r_{y'\underline{k}}=.00$, and $r_{y''\underline{k}}=.00$.

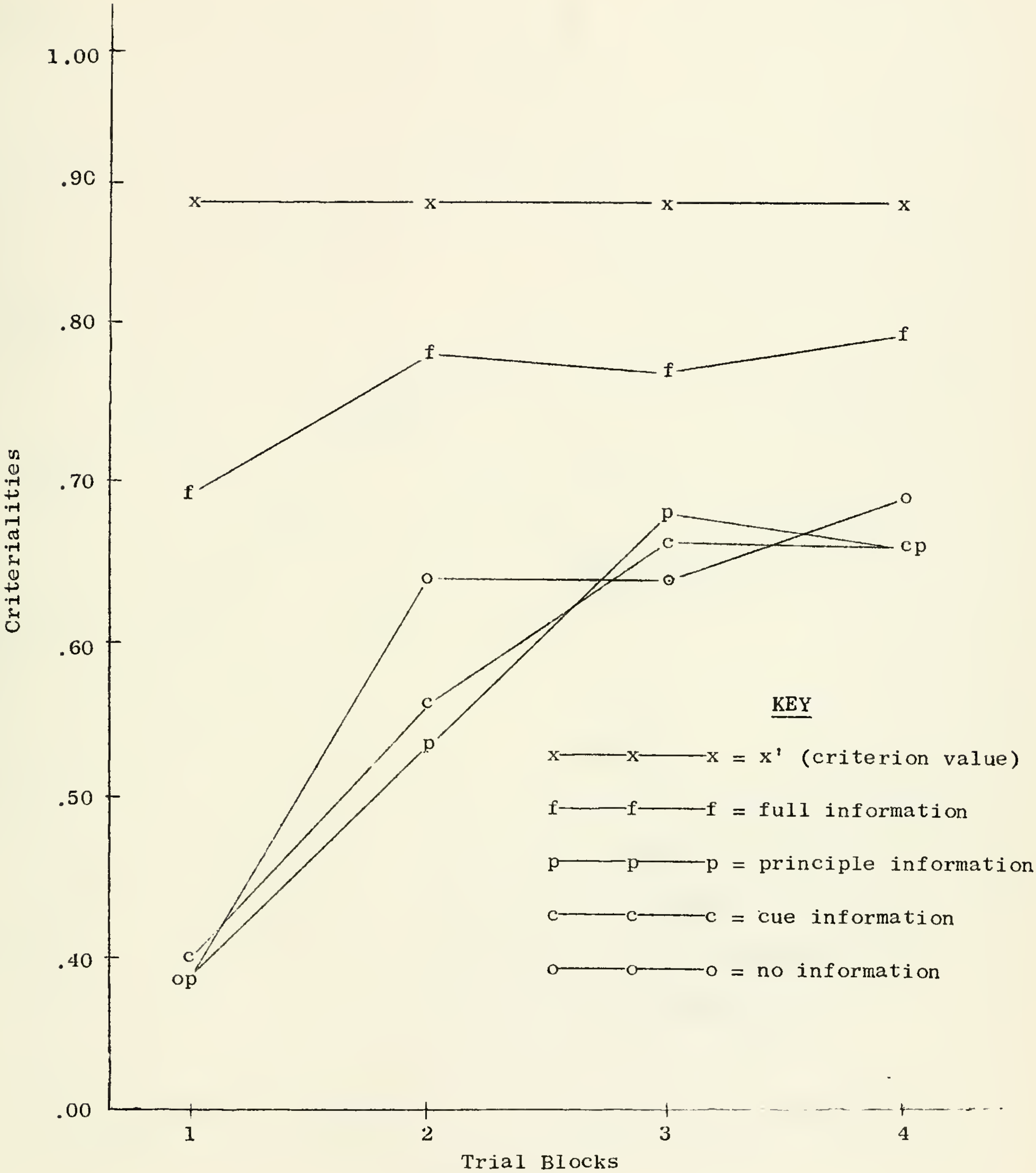


Figure 1. Mean criterialities of x' for each of the four experimental groups and criterion value of x'.

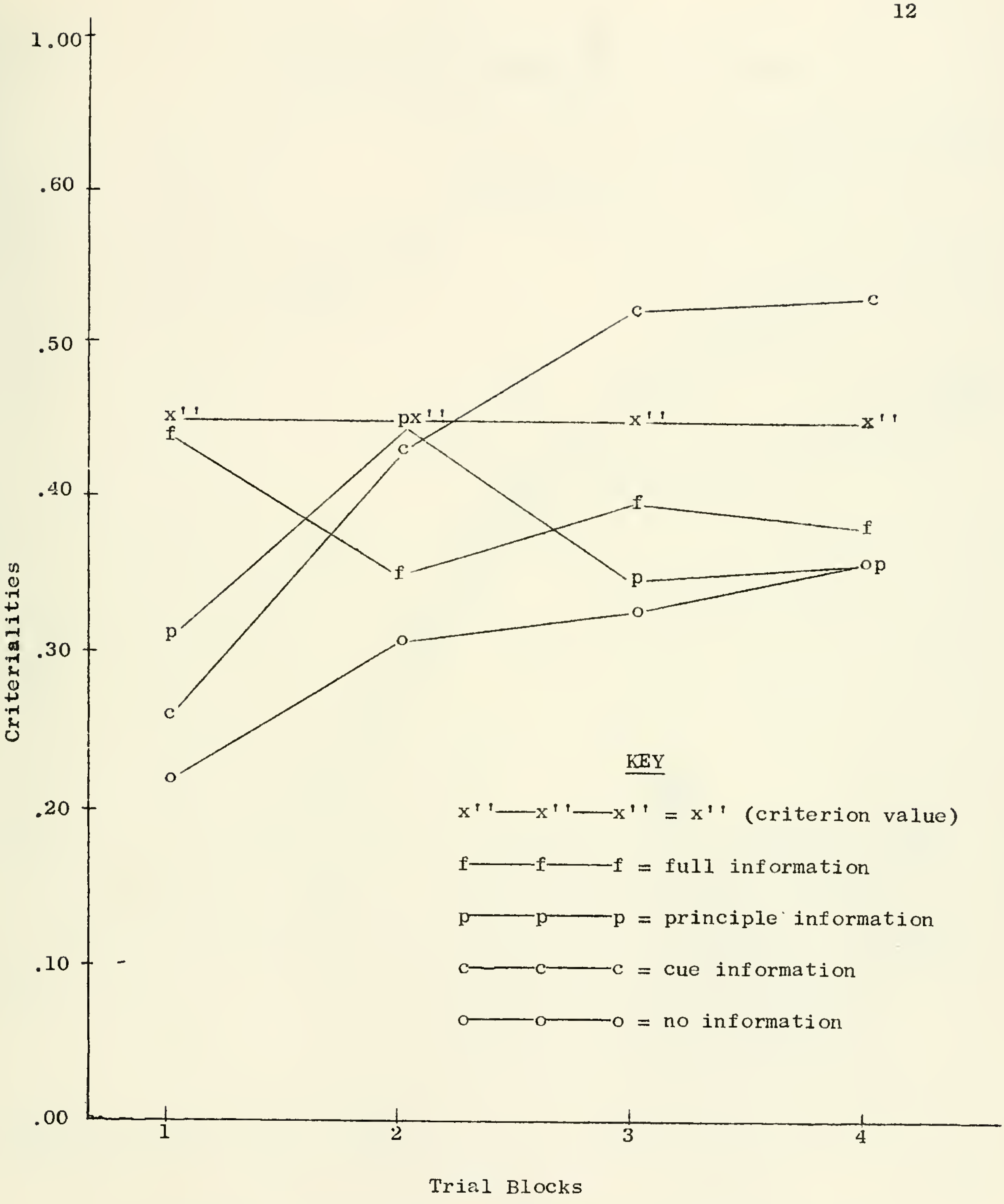


Figure 2. Mean criterialities of x'' for each of the four experimental groups and the criterion value of x''.

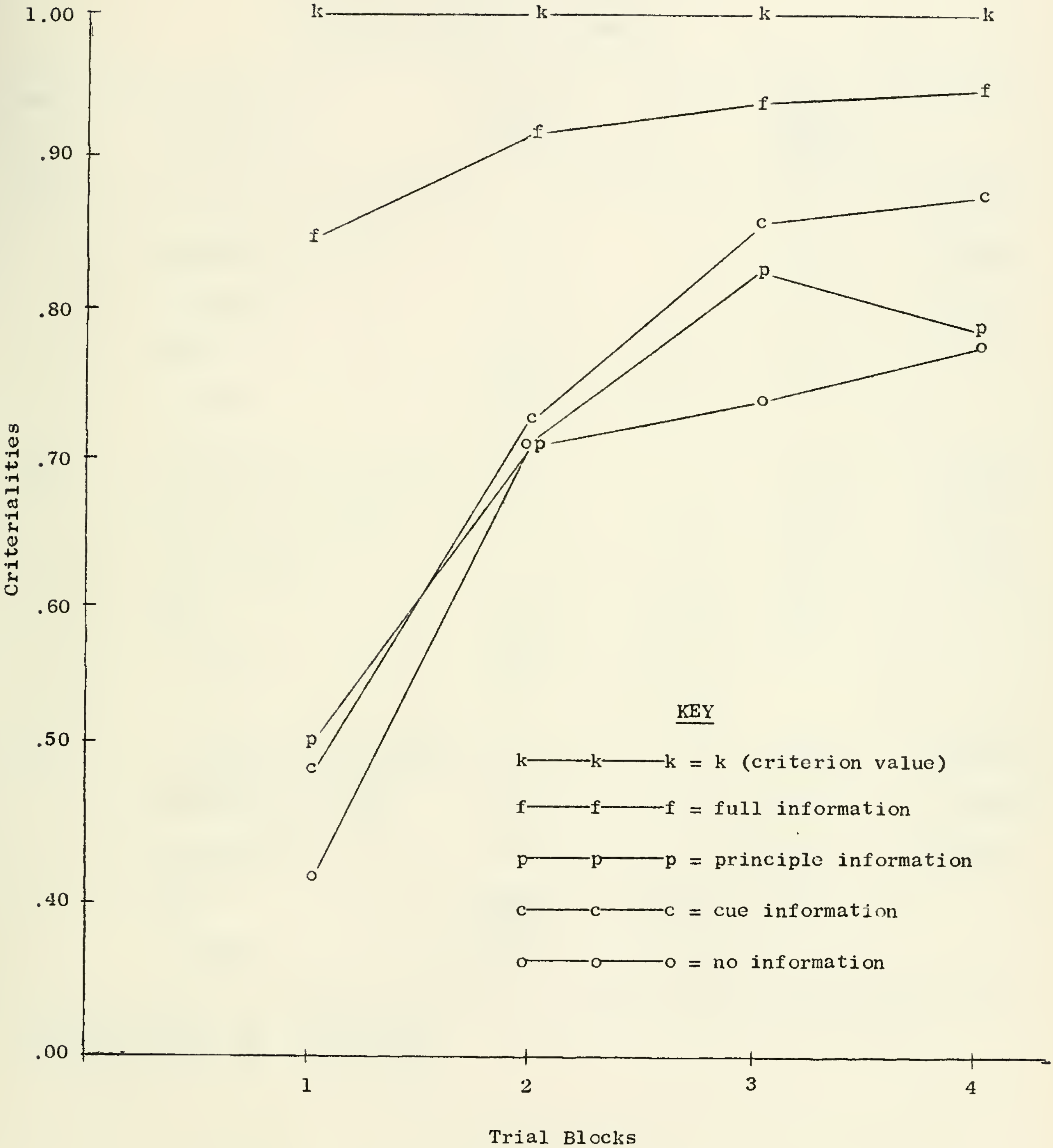


Figure 3. Mean criterialities of k for each of the four experimental groups and the criterion value of k.

Table 2

Analyses of Variance of x', x" and k for Trial Block 1

Variable	Source of variation[a]	Sum of squares	d.f.	Mean squares	F
x'	Knowledge of Principle	.58	1	.58	4.14*
	Knowledge of Cues	.82	1	.82	5.86*
	Interaction	.49	1	.49	3.50
	Within	6.56	48	.14	
	Total	8.45	51		
x"	Knowledge of Principle	.32	1	.32	8.00**
	Knowledge of Cues	.14	1	.14	3.50
	Interaction	.02	1	.02	.50
	Within	1.93	48	1.93	
	Total	2.41	51		
k	Knowledge of Principle	2.12	1	2.12	10.10***
	Knowledge of Cues	1.96	1	1.96	9.33***
	Interaction	1.20	1	1.20	5.71*
	Within	10.23	48	.21	
	Total	15.51	51		

*Significant at .05 level.
**Significant at .01 level.
***Significant at .005 level.

[a] The principle and full information groups were told the principle; cue and full information groups were told the number of the cues.

Table 3

Analyses of Variance of x', x"
and k for Trial Block 2

Variable	Source of variation[a]	Sum of squares	d.f.	Mean square	F
x'	Knowledge of Principle	1.23	1	.23	1.44
	Knowledge of Cues	.34	1	.34	2.12
	Interaction	1.03	1	1.03	6.44*
	Within	7.70	48	.16	
	Total	10.30	51		
x"	Knowledge of Principle	.02	1	.02	.33
	Knowledge of Cues	.00	1	.00	.00
	Interaction	.25	1	.25	4.17*
	Within	2.82	48	.06	
	Total	3.09	51		
k	Knowledge of Principle	1.41	1	1.41	4.08
	Knowledge of Cues	1.81	1	1.81	5.17*
	Interaction	1.30	1	1.30	3.71
	Within	16.73	48	.35	
	Total	21.25	51		

*Significant at the .05 level.

[a]The principle and full information groups were told the principle; the cue and full information groups were told the number of cues.

Table 4

Analyses of Variance of x', x"
and k for Trial Block 3

Variable	Source of variation[a]	Sum of squares	d.f.	Mean square	F
x'	Knowledge of Principle	.27	1	.27	1.08
	Knowledge of Cues	.13	1	.13	.05
	Interaction	.07	1	.07	.03
	Within	12.13	48	.25	
	Total	12.60	51		
x"	Knowledge of Principle	.05	1	.05	1.00
	Knowledge of Cues	.27	1	.27	5.49*
	Interaction	.10	1	.10	2.00
	Within	2.24	48	.05	
	Total	2.66	51		
k	Knowledge of Principle	1.39	1	1.39	2.24
	Knowledge of Cues	2.38	1	2.38	3.84
	Interaction	.11	1	.11	.18
	Within	29.75	48		
	Total	34.63	51		

*Significant at the .05 level.

[a]The principle and full information groups were told the principle; the cue and full information groups were told the number of cues.

Table 5

Analyses of Variance of x', x" and k for Trial Block 4

Variable	Source of variation[a]	Sum of squares	d.f.	Mean square	F
x'	Knowledge of Principle	.20	1	.20	.91
	Knowledge of Cues	.19	1	.19	.86
	Interaction	.36	1	.36	1.64
	Within	10.45	48	.22	
	Total	11.20	51		
x"	Knowledge of Principle	.13	1	.13	2.60
	Knowledge of Cues	.16	1	.16	3.20
	Interaction	.12	1	.12	2.40
	Within	2.56	48	.05	
	Total	2.97	51		
k	Knowledge of Principle	.76	1	.76	1.29
	Knowledge of Cues	4.05	1	4.05	6.86*
	Interaction	.67	1	.67	1.14
	Within	28.48	48	.59	
	Total	33.96	51		

*Significant at the .05 level.

[a]The principle and full information groups were told the principle; the cue and full information groups were told the number of cues.

Table 6

Results of Duncan's Multiple Range Tests for x', x" and *k* for Each Block of Trials

Variable	Block	Groups	Probability
x'	1	Full > No Information, Principle, Cues	< .01
	2	Full > Principle, Cues	< .05
	3	None	
	4	None	
x"	1	Full > No Information Full > Cues	< .005 < .05
	2	None	
	3	Cues > No Information, Principle	< .05
	4	Cues > Full, No Information, Principle	< .05
k	1	Full > No Information, Principle, Cues	< .005
	2	Full > No Information, Principle, Cues	< .01
	3	Full > No Information	< .05
	4	Full > No Information, Principle	< .05

An analysis of variance was run for three variables, x', x" and *k*, using data from the first block of trials (See Table 2). Using *S*'s *k* values as the dependent variable, knowledge of the principle and knowledge of the number of cues produced a significant effect on performance at the .005 level, and a significant interaction between these variables at the .05 level. These results can be explained by the superior performance of the full information group, which is significantly better than that of all other groups at the .005 level. Using the criterialities, it is clear that the full information group learned the most relevant cue (x') better (criteriality was .69) than any other group, and learned the less relevant cue (x") better (criteriality was .44) than any of the other groups (no information, .22; cue information, .26; principle information, .13).

In the second block of trials, knowledge of the number of cues continues to produce a significant effect at the .05 level. The superior performance of the full information group again accounts for the significant difference in performance. Since an F of 4.04 is necessary for significance at the .05 level, the F relating to knowledge of the principle (4.03) is slightly less than significant. The full information group's criteriality for x' is still significantly better (.78) than that of either the principle (.54) or cue groups (.56); interestingly enough, although better, it is not significantly better than the value for the no information

group (.64). For both relevant cues (x' and x") the interaction effect was significant on the second block of trials.

In the third block of trials, neither knowledge of the principle nor knowledge of the number of cues is significant. The performance of the full information group was significantly superior to the performance of the no information group, but it was not superior to that of the other groups. The mean criteriality of the cue group for x" (.52) is significantly higher than that of either the no information (.33) or principle group (.45). This difference is accounted for by the fact that the cue group tends to weight the relevant cues more evenly. In fact, the cue group's mean criteriality for x" is greater than the criterion criteriality for x".

In the fourth block of trials, it appears that having knowledge of the number of cues results in significantly superior performance at the .05 level. The full information group performed significantly better than either the no information or principle groups. The mean criteriality of $\underline{k}$ for the principle group dropped unexpectedly from .83 in block 3 to .79 in the fourth block. The mean criteriality for the cue group (.88) is approximately midway between that of the full information group (.95) and that of the no information (.78) and principle groups (.79). The mean criterialities for the latter two are almost identical. The mean criteriality for x" of the cue group (.53) is significantly greater than that of the other groups, including the full information

group (.38), which is the highest of the other groups. Even as late as the fourth block of trials, the cue group tends to weight relevant cues more evenly (x' criteriality is .66 and x" criteriality is .53). Table 7 gives the mean relative weights of relevant cues for each group.

The relative weighting of the two relevant cues can be found for each S by dividing his criteriality for x' by his criteriality for x". Although the criterion weighting is 2:1, rounding errors make the criterion weighting 1.98:1 in the computations. Since relative weightings are uninterpretable when the criterialities are small or negative, the mean weightings for groups were computed only for the last block of trials. Only those subjects whose criteriality for k surpassed .70 were included in the analysis; Ss whose criteriality was less than .70 were considered to be non-solvers since less than half of the variance of their responses could be accounted for by E's criterion values.

Although the standard deviation of the weights of those in the principle group is quite large, their mean weighting and that of the full information group are very close to the criterion weighting. The no information group tends to overweight the more relevant cue; however, the variance in this group is the largest of the four groups. The cue group tends to underweight x', and overweight x". Except for one S whose weighting was perfect, all Ss in this group had a weighting below 1.98:1.

[illegible] (.), [illegible] standard [illegible]. The [illegible] the [illegible] group [illegible] [illegible] 96 [illegible]. [illegible] which [illegible] weights of [illegible] in each group.

The [illegible] following [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] that [illegible] for [illegible]. [illegible] [illegible] [illegible] compounds [illegible] 2 [illegible] cases.

Although the standard deviation of the weights of those in the principle group is quite large, their mean weighting and that of the [illegible] information group are very close to the [illegible] weight. The [illegible] [illegible] this group [illegible] group tends [illegible] whose [illegible] weighting [illegible] 1.98.

Table 7

The Mean Relative Weight of Relevant Cues (x' and x"), and S.D. of Ratio for Each Group in Trial Block 4[a]

Group	$\underline{N}$[b]	Relative weights of x' and x"	S.D. of ratio
Full Information	11	2.04	.31
Principle Information	10	2.02	.83
Cue Information	9	1.32	.44
No Information	8	2.46	1.62

[a]Relative weights for trial blocks 1, 2 and 3 were uninterpretable (criterialities small or negative.)

[b]Only those subjects were included whose criteriality for $\underline{k}$ was .70 or greater.

Conclusions

The following conclusions about the experimental hypotheses were drawn:

1. The rank order of performance for the four groups is not as predicted, since the group given information about the number of relevant cues performed better than the group given the solution principle. The difference between these two groups, however, is not significant.

2. A knowledge of the principle (principle group) is not more beneficial than a knowledge of the number of cues (cue group), although a knowledge of the principle together with a knowledge of the cues (full information group) leads to better performance; however, the final level of performance for the full information group is not statistically better than that of the cue group.

3. When the principle and cue groups are compared, a knowledge of the cues does not seem to be more beneficial initially, nor does a knowledge of the principle seem to be more beneficial later in learning; however, the groups who knew the principle (full information and principle groups) did learn the relative importance of the two relevant cues better than the other groups.

Discussion

Since it is difficult to administer a task of this kind in large groups, unless the cooperation of Ss is guaranteed, the data reported here must be regarded as a pilot study. E must know that the Ss whose data are included are actually answering before the verbal feedback is given. In this experiment, the only criteria for eliminating Ss from the analyses was if they failed to answer a significant number of questions or if they gave the same answers for a long series of trials. Furthermore, since the instructions were exceedingly complicated and, at times, obscure to Ss, some clarifying verbal instructions were needed, though not included.

In their instructions, the cue information group was told that there were four cues, two relevant and two irrelevant, while the principle group was told the proper weighting of the two relevant cues. In order to make the cue group comparable to the full information group, they should have been told which were the two relevant cues so that the cue group would have had to learn only the proper weighting. If these more informative instructions had been given, the hypothesis that a knowledge of cues will be beneficial in the initial blocks of trials might have been born out.

The fact that the groups did not reach an asymptote of performance at the end of four blocks of trials suggests that more trials should be added. Four blocks of trials were convenient in that the whole experimental procedure was

accomplished in one hour; however, the time should be increased.

A further difficulty was revealed by the first question of the questionnaire, particularly as answered by the no information group. E defines k in terms of stimulus parameters with an algebraic equation; however, S may arrive at the same solution with a geometric model. He may, for example, draw a line between the two colored crosses, pick a point on this line, drop a perpendicular from it to the bottom of the square, and then use the bottom of the square as a scale for k. By so doing he will arrive at the correct answer, even though his model of the task differs from that employed by E. If this is an easier solution, as it appears to be, and if this is the usual method that would be employed by S if he approached the task with no information, then the principle information given in algebraic terms might make the task unnecessarily complicated. This might account for the failure of the principle group to reach a final level of performance that exceeded that of the no information group. It would also make generalizations from the data suspect.

Cronbach & Azuma (1961) reported that S need not have only one hypothesis about k; he may have a different hypothesis for different subsets of stimulus presentations. For example, S may have one solution if the two crosses are in the same column, another if they are one column apart, and so on. This is a factor to be examined, although this seems less

likely in the task presented here since it included ten response categories with exact numerical feedback, whereas Cronbach and Azuma's task used four response categories represented by four standard stimuli. On the other hand, this factor would not vitiate an analysis of overall group success in terms of the criteriality of k, but generalizations about the criterialities of x' and x" would have to be made cautiously since S might not be using them in the same combination on every trial.

Summary

An attempt was made to determine the relative importance of a knowledge of the solution principle or a knowledge of the number of relevant cues in a concept-attainment task. The overall success of the cue group was better (though not significantly) than that of the principle group; however, Ss in the principle group who reached an arbitrary level of success did learn the relative weighting of the two relevant cues better than did comparable Ss in the cue group.

REFERENCES

AZUMA, H. COMPARISON OF A CORRELATION WITH A PROBABILISTIC APPROACH TO CONCEPT LEARNING. UNPUBLISHED DOCTORAL DISSERTATION, UNIVER. OF ILL., 1960.

AZUMA, H. & CRONBACH, L.J. PERFORMANCE ON A CONCEPT-ATTAINMENT TASK WITH SCALED CUES. URBANA: UNIVER. OF ILL., 1961.

AZUMA, H. CRONBACH, L.J. PERFORMANCE ON A CONCEPT-ATTAINMENT TASK WITH SCALED CUES. II. PROBABILISTIC FEEDBACK. URBANA: UNIVER. OF ILL., 1962.

BRUNER, J.S., GOODNOW, JACQUELINE J. & AUSTIN, G.A. A STUDY OF THINKING. NEW YORK: WILEY, 1956.

BRUNSWIK, E. PERCEPTION AND THE REPRESENTATIVE DESIGN OF PSYCHOLOGICAL EXPERIMENTS. BERKELEY: UNIVER. OF CALIF. PRESS, 1956.

CRONBACH, L.J. & AZUMA, H. CAN WE TELL WHAT THE LEARNER IS THINKING FROM HIS BEHAVIOR. PAPER PRESENTED AT TRAINING RESEARCH CONFERENCE, WASH. UNIVER., ST. LOUIS, MO., 1961.

DUNCAN, D.B. MULTIPLE RANGE AND MULTIPLE F TESTS. BIOMETRICS 1955, 11, 1-42.

HOVLAND, C.I. A COMMUNICATION ANALYSIS OF CONCEPT LEARNING. PSYCHOL. REV., 1952, 59, 461-472.

HULL, C.L. QUANTITATIVE ASPECTS OF THE EVOLUTION OF CONCEPTS. PSYCHOL. MONOGR. 1920, NO. 123.

MCHALE, T.J. AND STOLUROW, L.M. CONCEPT FORMATION WITH METRICALLY MULTIVALUED CUES AND RESPONSE CATAGORIES. URBANA: UNIVER. OF ILL., TRAINING RESEARCH LAB., 1962.

SMEDSLUND, J. MULTIPLE PROBABILITY LEARNING. OSLO: AKADEMISK FORLAG, 1955.

SMOKE, K.L. NEGATIVE INSTANCES IN CONCEPT LEARNING. J. EXP. PSYCHOL. 1933, 16, 583-588.

I ~~D6P~~

TRAINING RESEARCH LABORATORY

Department of Psychology
University of Illinois
Bureau of Educational Research
8 Lincoln Hall
Urbana, Illinois

HOW CONSCIOUS IS TRANSFER OF A SPECIFIC RULE?

Thomas J. McHale

TECHNICAL REPORT NO. 6

Psychological and Educational Factors in
Transfer of Training
Phase II

August, 1965

U.S. Office of Education
Contract # 4-20-002

Principal Investigator
Lawrence M. Stolurow

TRAINING RESEARCH LABORATORY
Bureau of Educational Research
Department of Psychology
University of Illinois
Urbana, Illinois

HOW CONSCIOUS IS TRANSFER OF A SPECIFIC RULE?[1]

Thomas J. McHale[2]

TECHNICAL REPORT NO. 6

August, 1965

Psychological and Educational Factors in Transfer of Training
Phase II

Principal Investigator:
Lawrence M. Stolurow

Project Sponsor:
Educational Media Branch
U.S. Office of Education
Title VII
Contract # 4-20-002

---

[1] Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology in the Graduate College of the University of Illinois, 1965.

[2] Now at Milwaukee Institute of Technology

TABLE OF CONTENTS

Table of Contents (cont.)

Table of Contents (cont.)

LIST OF TABLES

List of Tables (cont.)

LIST OF FIGURES

How conscious is transfer of a specific rule?

Transfer can be defined as "the effect of a preceding activity upon the learning of a given task" (Osgood, 1953), or as a "change in ability to deal with situations not encountered during training" (Cronbach, 1963). According to the latter author, transferable outcomes include specific actions, specific facts, broad concepts and generalizations, techniques of analyzing situations, attitudes toward the subject or situation, and even attitudes towards oneself. Though Osgood's definition of transfer emphasizes learning whereas Cronbach's emphasizes ability, the difference between the two need not be more than this matter of emphasis. Learning and changes in ability can be viewed as two sides of the same coin. Anything learned, if it is transferable, is a change in ability, and most changes in ability are the result of some new transferable learning.

Problem

The present study is concerned with transfer resulting from the learning of a specific rule or processing formula. In the first task, the rule was learned as applied to one set of stimuli; in the second task, the same rule and other alternative rules are potential solutions. The question was: What are the specific mechanisms through which transfer occurs?

Transfer Mechanisms

Why transfer occurs when it does is an interesting problem. The idea that the transfer of specific rules is not automatic is certainly not a new one. Judd (1927) stated it long ago. In fact, the problem of how to teach for transfer is one which has constantly plagued educational psychologists.

Two contemporary psychologists have been offered suggestions:

> Transfer of a behavior pattern to a new situation is likely to occur whenever the person recognizes the new situation as similar to other situations for which the behavior has been appropriate. (Cronbach, 1963).

> Perhaps the most important single determinant of the amount of transfer that is possible, and that we usually can do something about, is the knowledge, on the part of the learner, that what he is learning can be transferred. (Deese, 1958).

Both suggest that the learner must know that what he is learning, or has learned, is applicable to new situations, and he must be able to recognize these new situations when they occur. The term "recognition" seems to imply some type of conscious process. However, no experiment seems to have deliberately investigated whether the mechanisms of transfer are conscious processes.

Awareness in learning. The role of awareness in human learning has recently been rekindled into an issue in learning theory. One group of studies has reported learning without awareness in verbal operant conditioning. These studies have been reviewed by Adams (1957), Krasner (1958), Salzinger (1959), and Eriksen (1960). As some of the reviewers have pointed out, the questions asked often seem inadequate and the criteria of awareness are sometimes vague and arbitrary. Critical studies of learning without awareness have appeared in the areas of verbal operant conditioning (Dulany, 1961, 1962; Spielberger, 1962), and motor operant conditioning (Paul, Eriksen & Humphreys, 1962). With more adequate questioning, better criteria of awareness, and recognition of correlated hypotheses, negative results have been reported.

The major theory resulting from the recent controversy about learning without awareness was proposed by Dulany (1962). This theory is one of propositional verbal control of behavior under selective reinforcement.

controversy, such labels have been offered as suggestions:

> Transfer [illegible] to put [illegible] a new situation is likely to occur [illegible] the [illegible] recognizes the new situation as similar [illegible] situations under which the behavior has been appropriate. (Grabbman, 1960).

> Perhaps the [illegible] determinant of the amount of transfer that is possible, and that we usually can do something about, [illegible] on the part of the learner, that what he is learning [illegible] be transferred. (Deese, 1958).

Both suggest that the learner must know that what he is learning, or has learned, is applicable to new situations, and he must be able to recognize those new situations when they occur. The term "recognition" seems to imply some type of conscious process. However, no experiment seems to have deliberately investigated whether the mechanisms of transfer [illegible] conscious processes.

Awareness in learning. The role of awareness in human learning has recently been rekindled into an issue in learning theory. One group of studies has reported learning without awareness in verbal operant conditioning. These studies have been reviewed by Adams (1957), Krasner (1958), Salzinger (1959), and Eriksen (1960). As some of the reviewers have pointed out, the questions asked often seem inadequate and the criteria of awareness are sometimes vague and arbitrary. Critical studies of learning without awareness have appeared in the areas of verbal operant conditioning (Dulany, 1961, 1962; Spielberger, 1962), and motor operant conditioning (Paul, Eriksen & Humphreys, 1962). With more adequate questioning, better criteria of awareness, and [illegible] of correlated hypotheses, negative results have been reported.

The major theory resulting from the recent controversy about learning without awareness was proposed by Dulany (1962). This theory is one of propositional verbal control of behavior and of selective reinforcement.

The name of the theory could be misleading. It does not state that words or covert speech control behavior. It states that intra-organismic processes or conscious states, called "hypotheses" and "intentions" control overt behavior. Though these conscious processes are assessed by verbalizations, in the learner's awareness they may be completely verbalized, only partially verbalized, or merely cognitive-neural. The criterion of awareness is the learner's ability to verbalize specific hypotheses or intentions when questioned. The theory is neutral with respect to idiosyncratic differences in conscious states.

Awareness in transfer. Three earlier studies suggest not only that there are conscious processes involved in some types of transfer, but that these processes are somewhat analogous to those proposed by Dulany in his theory. Ruger (1910) distinquished automatic from non-automatic transfer by asserting that the latter is dependent on an act of analysis or conscious control. While investigating the solution of a set of mechanical puzzles, he observed that "it is not the mere occurrence of a variation but its conscious continuance" which leads to quick solutions. In other words, a possible behavioral hypothesis (something that the learner considers testing or trying) does not become actually useful unless it is consciously pursued. Barker (1932) found that a hint to relate a second finger maze to an already learned maze produced faster solutions of the second maze. He explained the difference in terms of "factory," which he called "knowledge of a pattern relationship" or "a general idea which would serve as a control."

Partial transfer. Stolurow & McHale (1965b) found that transfer of rules is a highly complex phenomenon. Their results suggest that transfer is not automatic, that it can occur at various times, and that there are

individual differences in the manner in which it occurs. For example, if the rule learned in training is a complex one, it need not be transferred _in toto_. That is, a _S_ might transfer the cues alone, the principle alone, or anything else which is dissociable from the total rule. However two things in their study make the results difficult to interpret. Negative transfer occurred, which was difficult to handle within the context of the task. Since information about conscious transfer was obtained in post-experimental interviews, it is posssible that these after-the-fact reports might not accurately reflect the transfer process as it actually occurred. It would be better if they could be obtained simultaneously with the process itself.

Theory

Suggestions from three preceding experiments were combined into the tentative theory outlined in Figure 1 which was tested in the present experiment. This theory is analogous to that of Dulany (1962), and was stimulated by his thinking. When a rule has been learned, it exists in some form in the learner's memory. But this prior learning will have influence on learning in a new situation _only_ if a two-stage process occurs. This two-stage process includes both a transfer hypothesis (TH) and a transfer intention (TI). A "transfer hypothesis" can be described as the learner saying something like the following to himself: "I wonder if situation B is somehow or other related to situation A." But a transfer hypothesis alone is not sufficient for prior learning to have influence in a new situation. The learner might well decide against attempting to relate the new situation to an old one. Prior learning has an effect on the new situation only if a transfer intention accompanies the transfer hypothesis.

[illegible] differences in the [illegible] cure. For [illegible] what in training [illegible] not be [illegible] in [illegible] be primary [illegible] anything else which [illegible] Here [illegible] in their study [illegible] interpret. Here [illegible] which occurred, which [illegible] within the context [illegible]. Since information [illegible] obtained [illegible] experimental [illegible] after-the-fact [illegible] might not be accurate [illegible] as it actually occurred. It would be better [illegible] simultaneously with [illegible] process [illegible].

Theory

Suggestions from the previous [illegible] were combined into a tentative theory [illegible] tested in the present experiment. This theory is [illegible] Bulany (1962), and was formulated [illegible] it exists [illegible] some form in the learner [illegible] prior learning will have influence on learning [illegible] two-stage process [illegible]. This two-stage process [illegible] a transfer hypothesis (TH) and a transfer intention [illegible]. The "transfer hypothesis" can be described as the learner saying [illegible] himself: "[illegible] if situation B [illegible] situation A." But a transfer hypothesis alone is not [illegible] for prior learning to have influence in a new situation. The [illegible] first attempting to relate the new situation [illegible] prior learning [illegible] effect on the new situation only [illegible] transfer [illegible] transfer hypothesis.

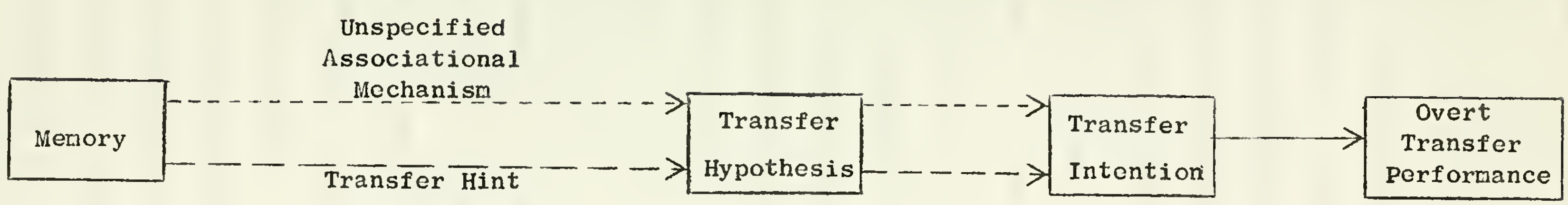


Fig. 1. Flow diagram of the tentative theory tested in this experiment.

A "transfer intention" is the actual conscious attempt to use old knowledge in a new situation. This "actual attempt to relate" may itself be temporary if the learner is not successful in finding a relationship. That is, the learner might decide to stop trying to relate the situations.

Ordinarily transfer hypotheses would seem to arise through some unspecified associative mechanism. If they are conscious processes; however, they also should be produced by instructions. That is, a transfer hint, similar to that given by Barker (1932), should produce both a transfer hypothesis (TH) and a transfer intention (TI). A higher proportion of Ss given such a hint should report transfer hypotheses and transfer intentions, and consequently groups given such a hint should solve the transfer task faster. Furthermore, within groups given a transfer hint, fewer Ss should report a transfer hypothesis without an accompanying transfer intention since the whole idea of transferring was given by E and not self-generated.

To test this theory, it is necessary to use some new dependent variables as measures of transfer. The customary measure, trails to criterion, is not the whole story. It is possible that a learner can attempt to relate prior learning to a new situation without being very successful. Yet his approach to the new situation would still show the influence of prior learning. If transfer is defined as any influence of specific prior learning on behavior in a new situation, trials to criterion is not a completely satisfactory measure. It must be supplemented with other measures of the learning process in the transfer situation such as verbal reports. Both trials to criterion and verbal reports were analyzed in this study.

The effect of verbalizations on performance can vary. Gagne & Smith (1962) found that performance was improved, while Stevenson & Weir (1963) and Toda (1962) found no difference in performance. Improvement was found

A "transfer intention" is the actual conscious attempt to use old knowledge in a new situation. This "actual attempt to relate" may itself be temporary if the learner is not successful in finding a relationship. That is, the learner may [illegible] to relate the situations.

Ordinarily transfer hypotheses would seem to arise through some unspecified associative mechanism. If they are conscious processes; however, they [illegible] should be produced by instructions. That is, a transfer hint, similar to that given by [illegible] (1963), should produce both a transfer hypothesis (TH) and a transfer intention (TI). A higher proportion of Ss given such a hint should [illegible] transfer hypotheses and transfer intentions, and consequently [illegible] Ss given a hint should solve the transfer task faster. Furthermore, [illegible] given a transfer hint, fewer Ss should report a transfer hypothesis [illegible] accompanying transfer intention since the whole [illegible] is given by E and not self-generated.

To test this theory it [illegible] necessary to use some new dependent variables as measures of transfer. The [illegible] measure, trails to criterion, is not the whole story. [illegible] possible that a learner can attempt to relate prior learning to a new situation [illegible] very successful. Yet his approach to the [illegible] show the influence of prior learning. If [illegible] the influence of specific prior learning on behavior in a new situation, trials to criterion is not a completely satisfactory measure. It [illegible] supplemented with other measures of the learning process in the [illegible] such as verbal reports. Both trials to criterion and verbal reports were analyzed in this study.

The effect of verbalizer [illegible] performance can vary. Gagne & Smith (1962) [illegible] that performance [illegible] while Stevenson & Weir (1963) and [illegible] (1962) found no [illegible]. Improvement was found with the "tower of Hanoi" problem, not with probability learning tasks. It is difficult to think of any a priori reason why thinking aloud should have any biasing effect on transfer in the present experiment in which all Ss including those in the control group were encouraged to think aloud and were required to give trial-by-trial hypotheses.

Method

Design. Forty-eight Ss were trained to criterion with each of the two training rules for a total of 96 experimental Ss. Twenty Ss in each training-rule group were given a transfer hint; the other 28 Ss in each group were not given this hint. Each subgroup had an equal number of males and females. Therefore this was a 2 x 2 x 2 (pretraining rule x sex x knowledge of the relatedness of the two tasks) design with 10 Ss in each transfer-hint cell, and 14 Ss in each no-transfer-hint cell. Besides the 96 experimental Ss, there were 32 control Ss with no relevant pretraining.

Subjects. About half of the Ss in the experiment participated as part of a course requirement for Introductory Psychology. The other half were fairly evenly divided between volunteers from other undergraduate psychology courses and paid volunteers. Three female Ss could not understand the rule for the training task, and refused to continue in the experiment on the grounds that numerical problems were too difficult for them. These three Ss were replaced. It is assumed that dropping these three Ss contributed to reducing group differences between males and females.

Task Model

The training and transfer tasks were generated from the same task model. The training task, originally developed by Azuma (1960), was used in a modified

form (Stolurow & McHale, 1964a). The transfer task, developed by Mattson (1963), was also modified to suit the purpose of the present study.

Training task. An example of a stimulus from the training task is given in Figure 2. The concept to be learned is the *k*-ness of each stimulus. The *k* value of a stimulus can be any whole number from 3 to 12, and so there are 10 possible answers. The circle and square can appear in any row and any column. The rows and columns are numbered from 1 to 4 from bottom to top and left to right respectively. Hash marks on the frame are included to eliminate *S*'s need to estimate the row and column values. In *E*'s task model, the *column* values of both the circle and square are relevant, whereas their row values are irrelevant.

Instructions - learning task. In the present experiment, *S*s were instructed to solve the training task in either one of two ways. Both of these solutions had been used by various *S*s in a previous experiment (Stolurow & McHale, 1965a). Subjects were given written instructions explaining the task model together with the particular rule they were to use; some examples showing the correct application of the rule were also included. The two rules were:

1) *k* = 2(column number of the circle) + 1 (column number of the square)

2) *k* = 3(column number of the circle) ± the number of columns from the circle to the square

+, if the circle is to the left of the square

-, if the circle is to the right of the square

These two rules are perfectly correlated. That is, they both generate the same numerical answer for any stimulus.

Transfer task. An example of a stimulus from the transfer task is given in Figure 3. Again the concept to be learned is the *k*-ness of each stimulus; *k* can be any whole number from 3 to 12. A circle, square, triangle, and

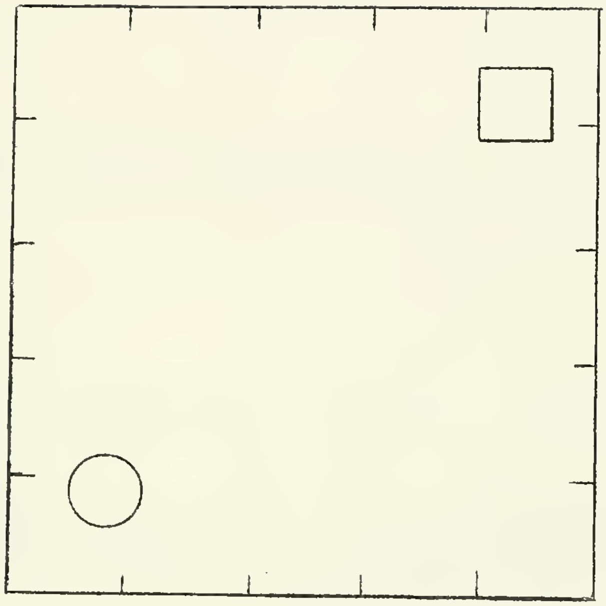


Fig. 2. A typical stimulus display of the training task.

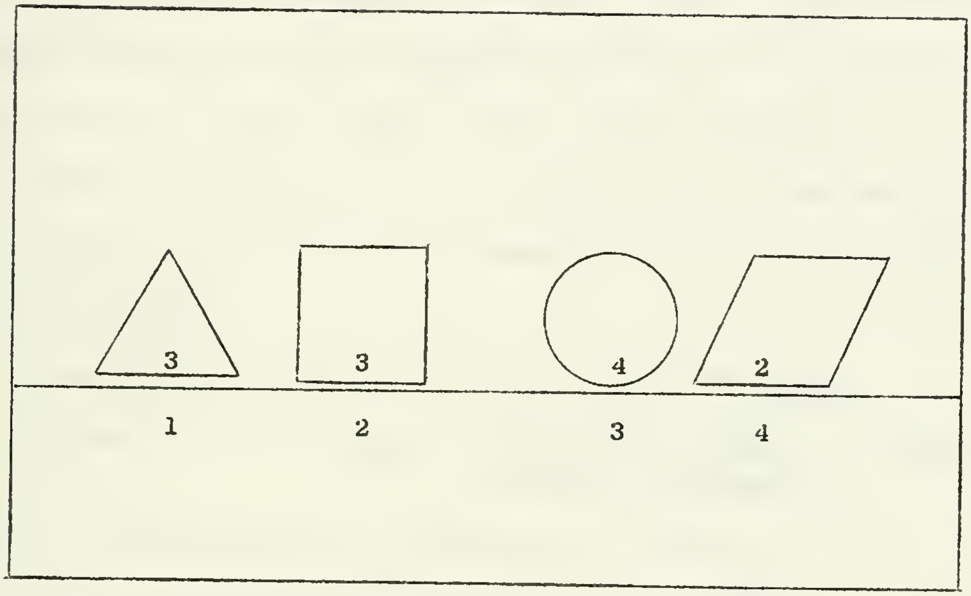


Fig. 3. A typical stimulus display of the transfer task.

rhombus appear on each stimulus card. Either a 1, 2, 3, or 4 appear in each figure; any number can appear in any figure. Each figure can appear in any of the four positions, counting from left to right. Either the numbers within the figures or the position of the figures can replace the row and column values of the training task. The figures are relevant in pairs. The position of the circle and square are relevant; the numbers within the triangle and rhombus are relevant. That is, the displaces were so designed that any of three rules would generate the same numerical answer for any one of them. Two of these rules are related to the position of the circle and square and are direct counterparts of the two training rules. The third is related to the numbers within the triangle and rhombus. The three rules are:

1) 2(position of the circle)+ 1(position of the square)

2) 3(position of the circle) ± the number of positions from the circle to the square

   +, if the circle is to the left of the square

   -, if the circle is to the right of the square

3) 2(number in the triangle) + 1(number in the rhombus)

These rules will be called 2CS, 3CS, and 2TR. It was impossible to design a counterpart using the numbers in the triangle and rhombus for 3CS. To do so would have eliminated all but a few of the possible cards.

Procedure

Training. Each S was run individually. After E explained that the experiment involved the investigation of S's manner of problem solving and that it would be explained afterwards, S was given a set of written instructions for either training Rule 1 or training Rule 2. When S finished read-

the instructions, he answered a set of questions about the relevant points in them. (The written instructions for both Rules 1 and 2 plus the set of questions for each can be found in Appendices A and B). If S answered any of the questions incorrectly, the correct answer was explained to him. The following instructions were then read by E:

> Here is a card with all of the possible values of k. I would like you to think aloud during this problem as you figure the answer for each card. This should not be too difficult since it's just a matter of applying the rule that was explained in the instructions. We'll continue until you give me 12 consecutive correct answers. So take your time and try to be as accurate as possible. I'll give you one card at a time, and I'll tell you the correct answer after you have given me yours. Here is some scrap paper if you think it is necessary. You may use it only for calculating purposes. That is, you may not keep a record of past figures or past correct answers. Are there any questions?
> You may refer to the instruction booklet at any time. Perhaps you would like to open it to the page on which the correct formula is given. Remember to think aloud as you work.

Both numerical answers and the formulas were recorded. The criterion of 12 consecutive correct answers (correct rule plus correct numerical answer) supplied evidence that S understood the rule. If S made a mistake related to a misunderstanding of the rule, E tried to correct his erroneous thinking since the purpose of the training task was to teach the rule and not to test how well the task instructions communicated. If S made a calculating mistake while using the correct rule, he was encouraged to strive for accuracy. The purpose of thinking aloud was twofold: 1) to detect any insights into other ways of solving the task, and 2) to get S accustomed to thinking aloud since this was a necessary part of the transfer task. No S discovered an alternate solution to the training task. When S attained criterion, all materials from the training task were removed and E immediately read the instructions for the transfer task.

Transfer task. E read the following instructions aloud while S read a second copy silently:

Here is a new problem. These are samples of the cards. Notice that there are four geometrical figures on each card: a rhombus, a circle, a square and a triangle. (Point to each figure on one card.) The geometrical figures can appear in any order on the cards. That is, each figure can be in either the first, second, third or fourth _position_, counting from left to right. (Describe the positions of one card). There is a _number_ inside of each geometrical figure; the numbers range from 1 to 4. Any number can appear in any geometrical figure. Both the _position_ and the _number_ of a particular figure will vary from card to card.

Again you have to decide how much _k_-ness each card has. _K_ can again be any whole number from 3 to 12. It can't be less than 3 or greater than 12, and it can't be a fraction or decimal, only a whole number. _K_ is our arbitrary name for some numerical concept. There is a correct formula by which _k_ is computed. The same identical formula applies to every card. It is your task to find this formula and use it correctly. You'll find it by giving me an answer (you'll have to guess on the first one), and then I'll tell you the correct answer and let you study the card to see how I might have obtained that particular answer. Are there any questions about the cards or the problem?

During this problem I'd like you to think aloud as much as possible. I'm interested in what goes on in your mind as you attempt to solve it, and I won't find out much about this unless you tell me. So try to tell me what you're considering, any conclusions you might come to, and any hunches or good guesses you might have. Since I'm not much of a mind-reader, don't think anything is too obvious to say aloud.

When I give you a card, you can answer as soon as you want, but if you don't answer within 30 seconds, I'll tell you to give an answer even if you have to guess. Think aloud during this time and tell me how you're getting the numerical answer you're testing. You may guess if you want, but please don't tell me you're guessing unless your answer has nothing to do with the position or numbers of the figure on the card. Then I'll tell you the correct answer, and you'll have 1 and 1/2 minutes to study the card. I'll warn you when there are 15 seconds left. If it takes you less than 1 and 1/2 minutes to decide on the formula you want to test next, just ask for the next card.

You may not use scrap paper at all during this problem. You'll have to do it all in your head. I know this is difficult, but some people would make better use of the scrap paper than others, and this would make the problem easier for them. Are there any questions about the procedure?

Remember that the position of each figure and the number within it changes from card to card. Not all of this information is necessarily used. Try to think aloud as much as possible. The experiment will not be a success unless you do this.

If any _S_ asked about information contained in the instructions, that information was re-explained. But if a _S_ asked about information not given

Here is [illegible] cards. [illegible] these are four geometrical figures [illegible], a square and a triangle. (Point [illegible].) The [illegible] figures can [illegible] each figure can be [illegible] the first, second, [illegible] counting from left [illegible] right. ([illegible]). There is a numb[illegible] [illegible] ranges from 1 [illegible] number [illegible] are. Both the position [illegible] the number [illegible] [illegible] to card.

Again you [illegible] much [illegible] any whole number [illegible] 16.1 [illegible] 12, and it can't be [illegible] section [illegible] arbitrary name for [illegible] meaningful concept. There [illegible] by which k is [illegible] same identical [illegible] every card. It is [illegible] task [illegible] this formula [illegible]. You'll find it by [illegible] (you'll [illegible] first one), and then [illegible] the correct [illegible] the card to see how [illegible] right [illegible] that [illegible] there any questions [illegible] [illegible] the [illegible]

During this problem [illegible] think aloud as [illegible] I'm interested in [illegible] [illegible] it, and I won't [illegible] that [illegible] as you [illegible] to tell me what [illegible] conclusions you [illegible] and [illegible] [illegible] [illegible] a mind-reader, [illegible]

When [illegible] give you a [illegible] you don't answer [illegible] even if you [illegible] how you're getting [illegible] if you want, but [illegible] unless [illegible] has [illegible] the card. [illegible] 1/2 minutes [illegible] left. If it [illegible] minutes to [illegible] formula you [illegible] the next card.

You may not [illegible] during this problem [illegible] to do it all [illegible]. I know this [illegible], but [illegible] would make both [illegible] the [illegible] there, [illegible] make the problem [illegible] any questions [illegible] procedure?

Remember that [illegible] the number [illegible] changes from [illegible] to [illegible] information is necessarily used. Try to think aloud [illegible]. The experiment [illegible] be a success unless you [illegible].

If any S asked [illegible] the instructions [illegible]

information was repeated [illegible] if [illegible] information [illegible] in the instructions, he was told: that is a good question to ask yourself. For control Ss all mention of a first problem was omitted. If S was in a transfer-hint group, the following instruction was read just before the task began:

> There's one hint I would like to give you. This problem with the second deck of cards is related to the first problem with the first deck of cards (show one). If you keep that first problem in mind, you should be able to solve this problem much more quickly.

This hint was typed on a card which was placed directly in front of S on the table.

Task hints. Task hints were given to all Ss in all groups three times during the problem, unless the problem had been solved by these times. The task hints were given so that all Ss would solve the problem in a reasonable amount of time. The three hints were given at trials 10, 20 and 30. They were:

At trial 10--Since this is a difficult problem, I will give you hints periodically. The first hint is this: there are only two figures which are relevant on each card. Two figures are relevant and you use them; two are irrelevant and you don't use them at all. The same two figures are relevant on each card.

At trial 20--I want to give you a final hint at this time. The correct formula for k is:

$$2a + b$$

That is, you multiply one relevant figure by two, and then add the second relevant figure. You'll have to discover what a and b are.

At trial 30--Position is entirely irrelevant in this problem. First, second, third and fourth positions don't count. It's the particular figures with particular shapes that count. With the approach you've been using you'll never solve the problem.

Since the task hints were given as soon as S gave his answer for cards 10, 20 and 30, he was able to study those cards in the light of them. These hints were also typed on cards and given to S to keep. The Ss were forced in th[illegible] tructions, [illegible] to[illegible]: th[illegible] good question t[illegible] ll.

[illegible] ter [illegible] first problem [illegible] mitte[illegible].

[illegible]ns-forming group, the [illegible] instruction was read just before the task [illegible]:

> [illegible] one hint I would like to give you. This prob[illegible] with the second deck of cards is related to [illegible] first probl[illegible] the first deck of cards (show one). If you [illegible] that first [illegible] in mind, [illegible] should be able [illegible] solve this [illegible].

[illegible] was typed [illegible] card which was placed directly [illegible] front of S on the [illegible].

Task hint[illegible]. Task hints were given to all Ss in all groups [illegible] times during the [illegible]. Unless the problem had been solved by th[illegible] time [illegible] the task hints w[illegible] on [illegible] that all Ss would solve the probl[illegible] reasonable [illegible] of time. The three hints were given at trials 10, 20 and [illegible]. They were:

At trial 1[illegible]--[illegible] this is a difficult problem, I will give y[illegible] particularly. The first hint is th[illegible]: there [illegible] figures which are relevant [illegible] each card. [illegible] are relevant and you use them; two are irrelevant [illegible] don't use them at all. The same two figures are relevant on each card.

At trial 20--I want to give you a [illegible] hint at this time. The correct formula for R is:

2a + b

That is, [illegible] multiply [illegible] relevant figure by two, and then add the second relevant figure. You'll have to discover what a and b are.

At trial [illegible]0--Position is entirely irrelevant in this problem[illegible] first, second, third and fourth positions don't [illegible] the particular figures with particular shape[illegible] with the approach you've been using you'll never solve the prob-

Since the [illegible] given as soon as S gave his an[illegible]

[illegible] study those cards in the light

hints were not [illegible] cards and given to S to keep. The [illegible] to use the hints in the sense that _E_ reminded them of the hint if they failed to use it on subsequent cards. The hint at trial 30 was included because of two _Ss_ who persisted in using formulas such as: 2(number in the first figure) + 1(number in the third figure). No task hint was given until trial 10 so that it was possible to investigate _Ss_' spontaneous approaches to the problem when they knew nothing about the task model. The criterion for the transfer task was four consecutive correct responses (correct formula plus correct numerical answer).

If an _S_ made a calculating mistake while thinking aloud in the 1 and 1/2 minutes he had to study the card after feedback had been given, _E_ asked him to check that formula again. Because of this technique, there might have been a small facilitating effect which was indirectly an effect related to verbalization. This type of monitoring is impossible if _Ss_ do not think aloud.

The following questions were asked after _S_ had attained criterion:

1) Did you notice any other formula that would have worked besides the one you have been using?

2) Did you think of the first problem with the first deck of cards while you were attempting to solve this second one? If so, what did you think of, and approximately when? (_E_ made it clear to _S_ that the question referred to the time before solution--not when the correct solution was discovered or when this question was asked.)

The purpose of the first question was to detect any other correct formula that _S_ might have noticed but not mentioned in thinking aloud or responding to each card. The purpose of the second question was to determine how well the experimental technique of thinking aloud detected transfer when it occurred. The information was used to supplement that obtained by _S_'s spontaneous verbalizations during the task itself.

Transfer elements and the task model. The experiment was so designed that any transfer which occurred was positive. Both pretraining rules can

be broken down into transfer elements. This logical analysis into transfer elements is in no way meant to suggest that the pretraining formula cannot be transferred as a unit. But these elements are logically dissociable from the complete rule, and on the basis of pilot work, Ss do dissociate parts of the complete formula when they transfer. The transfer elements from each pretraining rule are presented in Table 1.

Horizontal position means using the position of a figure (rather than the number inside it) to determine its value. Relative horizontal position means using the position of one figure in relation to another in order to determine whether to add or subtract a constant and how large the constant should be. A third use of position is possible in the transfer task. This third use is exemplified by the formula: add the numbers in the first three figures. The numbers within the figures are added, but no consideration is given to the particular figures which appear in these positions. This use of position is unrelated to the task model, and so it is called irrelevant position. The third task hint was given to eliminate the use of irrelevant position.

Control task. For a warm-up task and to introduce them to thinking aloud, the Control Ss solved items from Raven's Progressive Matrices Test (1938) for 10 minutes before the transfer task was begun. Ten minutes was the best estimate of the amount of time required for the whole pretraining of experimental Ss. Since pretraining Rule 2 seems to be more difficult to grasp and use than pretraining Rule 1, half of the control Ss were given "easy" items; half were given "difficult" items. The easy items were subtests A, B and C. The difficult items were subtests D and E. The control group was split this way to see whether an "easy" or "difficult" set for the second problem would develop. Control Ss were encouraged to think aloud be ... a ... with ... that ...

placed ... which ... pretraining ... result ...

... logical ...

... complete ... the ... do ...

the complete ... The ... element ...

pretraining ...

Shift ... test ... of a figure (rather than

the number ... relative horizontal ...

... in relation to another in ...

... pattern ... the ...

... in the transfer task ...

... split ... the ... in the fi...

... but no consider...

given ... figures which ... in those positions, ...

of positions ... is called ...

position, ... indicate the use ...

position.

Control I. ... task ... into ... than ...

alone, the ... items from Raven's progressive Matrices ...

(1938) for 10 minutes ... the transfer task was begun. The ... minutes ...

the best estimate of the amount of time required for the ... pretraining ...

experimental Ss. Since pretraining rule 3 seems to be more difficult ...

... used then pretraining rule 1, half ... the control Ss were given

"easy" ... ; half were given "difficult" items. The only items were sub-

tests ... C. The difficult items were subtests D and E. The control

group was ... this way to see whether an "easy" or "difficult" set ... the

second ... Control Ss were encouraged to think aloud

Table 1

Transfer Elements from Each
Pretraining Rule

Rule 1	Rule 2
1) Circle and square	1) Circle and square
2) 2 to 1 weighting	2) 3 to $\pm$ 1, $\pm$ 2, $\pm$ 3 weighting
3) Use of only two figures	3) Use of only two figures
4) Horizontal position	4) Horizontal position
	5) Relative horizontal position

during the warm-up task. With prompts if necessary, they were successful with each item, but E would occasionally mention how "easy" or "difficult" the items were. Since there were no time limits on particular items and E allowed S to proceed more or less at his own pace, the data for this warm-up task were not analyzed. The procedure with control Ss was identical to that of experimental Ss for the transfer task.

Experimental Hypotheses

Based on the tentative theory being tested in this study, the following predictions about the results were made. Predictions were made in terms of three dependent variables: trials to criterion, type of solution, and various verbal process measures.

Trials to criterion. 1) The effect of the pretraining rule plus its interaction with the transfer hint will be slight, if there is any. 2) The transfer-hint groups will solve faster than the no-transfer-hint groups. 3) There will be no difference between Ss without a transfer intention (those who do not consciously transfer) and control Ss.

Types of solution rules. 1) There should be a significant difference between groups trained with Rule 1 and Rule 2, with 3CS solutions appearing only among the latter group. 2) Analyzing the two training-rule groups separately, there should be no difference between Ss given a transfer hint and Ss who were not, but who report a transfer intention. 3) There should be no difference between Ss not reporting a transfer intention and control Ss.

Process measures. The process measures of most interest are those related to the task model such as the number of figures used, the particular figures used, types of weighting used, and whether horizontal position and relative horizontal position are used. 1) Among Ss reporting a transfer

intention, there should be no difference between those given a transfer hint and those not given this hint if they were trained with the same rule.

2) There should be no difference between Ss not reporting a transfer intention and control Ss, and the learning processes of neither group should be highly related to E's task model.

Results

Training Task

Because of E's experimental procedure of trying to clarify difficulties whenever they occured, the trials-to-criterion measure for the training task was not analyzed.

Rule difficulty. Of the 48 Ss taught Rule 1, only 7 made a total of 8 errors; of the 48 Ss taught Rule 2, only 17 made a total of 24 errors. The difference in number of Ss making errors with the two rules was signigicant ($X^2 = 4.50$, $p < .05$). Apparently the second rule was more difficult to learn and use.

Analysis of variance of two time measures for the training task support the conclusion that Rule 2 was the more difficult. The two time measures were: (1) time required to read the instructions and fill out the set of questions for them; and (2) time required to attain criterion--12 consecutive correct answers using the correct formula. Both were analyzed in a 2 x 2 (training rule x sex) analysis of variance. Means and standard deviations for each analysis are presented in Tables 2 and 4, respectively. The analysis are summarized in Tables 3 and 5, respectively. Type of rule was significant in both analyses; it took longer both to read the instructions and answer questions, and to attain criterion for Rule 2. The fact that Rule 2 seemed more difficult than Rule 1 supports the idea and [illegible] between [illegible] transfer [illegible]

in [illegible] realize with [illegible]

(2) [illegible] reported [illegible]

[illegible]

[illegible]

Training Task

[illegible] according to clarify [illegible]

[illegible] measure for the [illegible]

[illegible]

This difficulty [illegible]

[illegible] of the [illegible] found [illegible],

[illegible] two rate [illegible]

[illegible] presently the second [illegible] with

[illegible]

[illegible] training test

support [illegible] that Rule [illegible] result. The [illegible]

[illegible]: (1) [illegible] the instructions [illegible]

the [illegible] required to attain [illegible]

12 [illegible] formula. [illegible]

analyzed [illegible] analysis of variance. Means and

standard deviations [illegible] presented in Tables 2 and [illegible],

respectively. The [illegible] 5, respectively.

Type of rule was significant [illegible] long [illegible]

the instructions [illegible] criterion for [illegible].

The fact that Rule [illegible] supports [illegible]

Table 2

Means and Standard Deviations of Time Required to Read Instructions and Answer Questions for the Training Task

Groups	Mean	n	Standard Deviation
Rule 1			
Males	5.6[a]	24	1.06
Females	6.6	24	1.69
Rule 2			
Males	7.3	24	1.83
Females	8.4	24	1.33

[a]Means are correct to the nearest tenth of a minute

Table 3

Analysis of Variance of Time Required
to Read Instructions and Answer Questions
for the Training Task

Source	df	MS	F
Rule	1	71.24	31.4***
Sex	1	27.62	12.2***
Rule x sex	1	0.02	00.0
Within	92	2.27	
Total	95		

***Significant at .001 level

Table 4

Means and Standard Deviations of Time Required to Attain Criterion in the Training Task

Groups	Mean	n	S.D.
Rule 1			
Males	1.9[a]	24	0.75
Females	2.3	24	1.05
Rule 2			
Males	2.9	24	1.36
Females	3.3	24	1.41

[a]Means are correct to the nearest tenth of a minute

Table 5

Analysis of Variance of Time Required
to Attain Criterion in the Training Task

Source	df	MS	F
Rule	1	24.91	17.8***
Sex	1	3.88	2.8
Rule x sex	1	0.03	0.0
Within	92	1.40	
Total	95		

***Significant at the .001 level

of dividing control Ss into "easy" and "difficult" warm-up groups, since it is possible that a set for difficulty level did develop.

Transfer Task

Sex differences. Of the 24 Ss who did make mistakes 15 were males. The difference in the number of each sex making mistakes was not significant ($\chi^2 = 1.39$, $p > .20$). It took the females significantly longer to read the instructions, but not to attain criterion with Rule 2.

The results of the transfer task were analyzed in terms of three dependent variables: trials to criterion, type of solution, and various measures of the learning process. To test the experimental hypotheses, Ss were categorized into those who reported transfer hypotheses and transfer intentions and those who did not. These categorizations were based both on the spontaneous verbalizations during the task and the post-experimental interviews. For example, if an S did not spontaneously verbalize a transfer hypothesis (TH) or a transfer intention (TI) during the task but did so during the interview afterwards, he was still categorized as positive in that particular category. This supplementary use of the interviews seemed justified for two reasons: (1) the Ss were not trained to verbalize and they could not be instructed to verbalize transfer information; and (2) the proportion of Ss classified as positive TH or TI solely on the basis of the post-experimental interview was not different for the transfer-hint and no-transfer-hint groups. Given the assumption that Ss in the transfer-hint group did consciously transfer since they had the transfer hint on a card in front of them, then the proportion of Ss in that group who spontaneously verbalized transfer information during the task can be used as a base rate to justify the use of such information in categorizing no-transfer-hint Ss.

[illegible] "easy," [illegible] "difficult" were [illegible]
[illegible] ed the tasks.

[illegible]

[illegible] Th[illegible]
[illegible] nonb[illegible] ches [illegible]
[illegible]) [illegible] rticularly [illegible]
[illegible]

The results of the [illegible] three dependent variables: trials [illegible], and various measures of the learning pr[illegible] hypotheses, [illegible] were categorized into [illegible] hypotheses and transfer intentions and those who did [illegible]. These categ[illegible] on the spontaneous verbalization [illegible] intervie[illegible] interviews. For example, if [illegible] hypothesis (TH) or a transfer intention (TI) during [illegible] ing the interview afterwards, [illegible] categorized [illegible] positive in that particular category. This [illegible] seemed justified for two reasons: (1) the [illegible] to verbalize and they could not be instructed [illegible] transfer intention; and (2) the proportion of $\underline{S}$s classified as positive TH or TI [illegible] on the basis of the post-experimental interview was not [illegible] or the transfer-hint [illegible] transfer-hint groups. Given the assumption that [illegible] in the transfer group did conscious[illegible] transfer since they [illegible] the transfer hint on a [illegible] in front of them, then the proportion of $\underline{S}$s in that group who spontaneously verbalized transfer informati[illegible] during the task can be [illegible] as a [illegible] rate to justify the [illegible] information in categorizing [illegible] transfer-hint $\underline{S}$s.

In the transfer-hint group, 33 Ss reported a TH during the task, 7 Ss afterwards; in the no-transfer-hint group, 20 Ss during the task, and 7 Ss afterwards. This difference was not significant ($X^2 = 0.30$). In the transfer-hint group, 33 Ss reported a TI during the task, 7 Ss afterwards; in the no-transfer-hint group, 18 Ss during the task, and 6 Ss afterwards. This difference was not significant ($X^2 = 0.15$).

No matter whether they were given a transfer hint or not, Ss were classified as positive TH or positive TI if their reports fit the stated descriptions. In positive TH, at some time or other during the transfer task, the S wondered whether the second task might not be related to the first task. In positive TI, the S consciously attempted to use what was learned in the first task while solving the second task. The positive TI category is a gross classification since it ignores finer discriminations such as how and when and how long an S attempted to relate the two tasks. Two Ss who were classified as positive TI reported a temporary TI which was abandoned when unsuccessful. Among the other positive TI Ss, there were obvious differences in the time and manner in which they related the two tasks, but no meaningful categories could be found which would take this information into account.

The number of Ss in each experimental group classified as positive or negative for both TH and TI is given in Table 6. Of the transfer-hint Ss, all 40 reported both a positive TH and a positive TI. Of the 56 no-transfer-hint Ss, 27 reported a positive TH, 24 a positive TI. There was a significant difference between these two groups in the number reporting a positive TH ($X^2 = 27.3$, $p < .001$) and a positive TI ($X^2 = 31.8$, $p < .001$). This difference was predicted. Furthermore, no-transfer-hint S reported a positive TH without a corresponding positive TI, whereas there were three no-transfer-hint Ss who did. This higher degree of correspondence for transfer-hint Ss was also predicted.

the transfer-hint group, 78 $\underline{S}s$ [illegible] a TH during the [illegible] $\underline{S}s$ [illegible] in the non-transfer [illegible] during the [illegible] $\underline{S}s$ [illegible]. This [illegible] of significance (χ^2 [illegible]) [illegible] 31 $\underline{S}s$ [illegible] TH [illegible] 18 $\underline{S}$ [illegible]

[illegible] whether the [illegible] $\underline{S}s$ were classified as [illegible]. In positive [illegible] transfer task [illegible] whether [illegible] might not [illegible] task. In positive TH, the [illegible] to [illegible] the test task while solv[ing] [illegible] positive TH [illegible] classification since it ignores [illegible] as [illegible] when and [illegible] $\underline{S}$ [illegible] classified as positive TH reported [illegible] unsuccessful. Among the other positive [illegible] differences [illegible] which they [illegible] the [illegible] no meaningful [illegible] could be found which would take this [illegible] into account.

The number of $\underline{S}s$ in each experimental [illegible] negative for both TH and [illegible] is given [illegible]. [illegible] $\underline{S}s$, all 40 reported both a positive TH and [illegible] $\underline{S}s$, [illegible] 27 reported [illegible] TH [illegible]. There was a significant difference between these [illegible] number [illegible] positive TH ($\chi^2 = 27.7$, $\underline{p} < .001$) [illegible] $\underline{p}$ [illegible]. This difference was predicted. Furth[ermore] [illegible] $\underline{S}$ reported a positive TH without corresponding pos[itive] [illegible] were three non-transfer [illegible]. This [illegible] for transfer-hint $\underline{S}s$ [illegible] predicted.

Table 6

Number of Ss Classified as Positive and Negative TH and TI in Each Experimental Group

Group	TH		TI	
	Pos.	Neg.	Pos.	Neg.
Rule 1				
Transfer hint				
Males	10	0	10	0
Females	10	0	10	0
No transfer hint				
Males	6	8	6	8
Females	6	8	5	9
Rule 2				
Transfer hint				
Males	10	0	10	0
Females	10	0	10	0
No transfer hint				
Males	9	5	8	6
Females	6	8	5	9

Trials to criterion. The dependent variable for these analyses was the number of trials up to, but not including, the four criterion trials (trials on which correct answers and correct formulas were given). An analysis of variance was performed on trials to criterion for the basic experimental groups. Means, standard deviations, ranges and cell frequencies for this analysis are given in Table 7. A summary of the analysis is given in Table 8. The only significant main effect was transfer hint. No other main effect and no interaction was significant.

Pooling positive TH across training rule and sex, the mean and standard deviation for transfer-hint Ss were 5.85 and 6.1; for no-transfer-hint Ss, they were 8.41 and 7.7. For positive TI, the mean and standard deviation for transfer-hint Ss were 5.85 and 6.1; for no-transfer-hint Ss, they were 6.92 and 6.8. There was no significant difference between these transfer-hint and no-transfer-hint Ss for either TH (t = 1.55) or TI (t = 0.63). This lack of difference was predicted. Means, standard deviations, ranges and cell frequencies for these pooled TH and TI classifications are given in Tables 9 and 10.

Granted that the transfer hint did have a significant effect, the question still remains whether it had this effect through the mediation of conscious cognitive processes. To answer this question, it is necessary to examine the relationships of sex, training rule, transfer hint, TH, TI, and trials to criterion. The matrix of intercorrelations of these variables is given in Table 11. Since there was so significant difference between transfer-hint and no-transfer-hint Ss categorized as positive TH and TI, these Ss were pooled for all correlations. All correlations in this matrix are phi-coefficients except for correlations with trials to criterion, which are point biserials. Only the correlations between the transfer hint, TH, [illegible] was transfer-hint, [illegible]

[illegible] was significant.

[illegible] criterion rule [illegible] sex, the mean and standard deviation for transfer-[illegible] were [illegible] and [illegible]; for non-transfer-hint Ss, they were [illegible] and [illegible]. [illegible] criterion II, the mean and standard deviation [illegible] transfer-hint Ss [illegible]; for non-transfer-hint Ss, they were [illegible]. [illegible] no significant difference between [illegible] non-transfer-[illegible] Ss [illegible] criterion I [illegible] = 1.35) or II ($\underline{t}$ = [illegible]). The [illegible] means, standard deviations, ranges [illegible] the I and II classifications are given in Tables [illegible].

Granted that [illegible] have a significant effect, the question still remains [illegible] it has this effect through the mediation of consistent [illegible] codes. To answer the question, it is necessary to examine the relationships [illegible] sex, [illegible] rule, transfer hint, [illegible] and [illegible] as a criterion. The matrix [illegible] correlations of these [illegible] is given in Table [illegible]. [illegible] shows that [illegible] significant difference between transfer-hint and non-transfer-hint Ss [illegible] as positive TH and TI, [illegible] Ss were pooled for all [illegible] correlations in this matrix [illegible] phi-coefficients, except for [illegible] with trial [illegible] criterion, which [illegible] point biserial. Only the correlation [illegible] and the transfer hint, TH,

Table 7

Means, Standard Deviations, Ranges, and Cell Frequencies of Trials to Criterion for the Basic Experimental Groups

Group	Means	n	S.D.	Range
Rule 1				
Transfer hint				
Male	2.6	10	3.3	0 - 10
Female	7.2	10	7.7	0 - 20
No transfer hint				
Male	12.4	14	8.2	0 - 26
Female	12.8	14	6.4	5 - 24
Rule 2				
Transfer hint				
Male	7.1	10	5.8	2 - 21
Female	6.1	10	6.6	0 - 20
No transfer hint				
Male	11.4	14	9.7	0 - 24
Female	15.6	14	8.3	0 - 30

Table 8

Analysis of Variance of Trials to Criterion for the Basic Experimental Groups

Source	df	MS	F
Training Rule	1	37.5	0.68
Sex	1	104.2	1.89
Transfer Hint	1	1,238.6	22.44***
Rule x Sex	1	0.1	0.00
Rule x Hint	1	3.4	0.06
Sex x Hint	1	1.3	0.02
Rule x Sex x Hint	1	126.7	2.30
Within	88	55.2	
Total	95		

***Significant at the .001 level

Table 9

Means, Standard Deviations, Ranges and Cell Frequencies of Trials to Criterion for Positive and Negative Transfer Hypothesis Groups

Group	Means	n	S.D.	Range
<u>Rule 1</u>				
<u>Positive TH</u>				
Male	4.3	16	4.8	0 - 13
Female	8.3	16	6.9	0 - 20
<u>Negative TH</u>				
Male	16.2	8	7.9	1 - 26
Female	14.8	8	6.6	8 - 24
<u>Rule 2</u>				
<u>Positive TH</u>				
Male	7.0	19	7.3	0 - 21
Female	8.1	16	7.5	0 - 21
<u>Negative TH</u>				
Male	19.4	5	4.5	12 - 24
Female	18.8	8	7.1	8 - 30

Table 10

Means, Standard Deviations, Ranges, and Cell Frequencies of Trials to Criterion for Positive and Negative Transfer Intention Groups

Group	Means	n	S.D.	Range
Rule 1				
Positive TI				
Male	4.3	16	4.8	0 - 13
Female	7.6	15	6.5	0 - 20
Negative TI				
Male	16.2	8	7.9	1 - 26
Female	15.2	9	6.3	8 - 24
Rule 2				
Positive TI				
Male	6.3	18	6.6	0 - 21
Female	7.2	15	6.9	0 - 21
Negative TI				
Male	19.7	6	4.1	12 - 24
Female	19.5	9	6.7	8 - 30

Table 11

Intercorrelation Matrix of All Variables with Each Other and with Trials to Criterion

	Sex	Rule	Hint	TH	TI	Trials to Criterion
Sex		.00	.00	.07	.09	.13
Rule			.00	-.07	-.04	.08
Hint				.56*	.60*	.44*
TH					.93*	.57*
TI						.64*

*Significant at the .01 level

TI and trials to criterion were significant. Notice the following points:

1) The order of magnitude of the significant correlations with the criterion is as predicted: TI (.64) is higher than TH (.57) which is higher than the transfer hint (.44).

2) Since of the 67 Ss who reported a positive TH only three did not report a positive TI, the correlation between TH and TI is very high (.93).

3) Since the three Ss who reported a positive TH without a corresponding positive TI were all in the no-transfer-hint group, the correlation between transfer hint and TI (.60) is higher than between transfer hint and TH (.56).

Since the multiple correlation of all variables with the criterion is .66, in a purely predictive sense the use of TI alone is almost as good as using all possible variables.

In order to test for intrinsic relationships with the criterion, each variable was correlated with the criterion with all other variables partialled out. These partial correlations were: sex (.10), training rule (.13), transfer hint (.10), TH (-.08), and TI (.32). Since only the last correaltion is significant, only TI is intrinsically related to the criterion. This correlation is greatly reduced because of the high degree of correspondence between TH and TI. The correlation of TI and the criterion with all variables except TH partialled out is .52. It seems improbable that "deciding not to pursue a transfer hypothesis" would ever be a common phenomenon given this experimental situation. Introduced to a novel situation, Ss are attempting to solve a complex problem with no leads to follow except what they have learned in the first task.

An analysis of variance (difficulty level of warmup task x sex) of trials to criterion for the control group was also performed. The means, standard deviations and ranges for subgroups are given in Table 12; the summary of the analysis is given in Table 13. Since there were no significant

TI on criteria [illegible]:

1) The [illegible] correlations with the criterion [illegible] (b) [illegible] higher than [illegible] (.[illegible]) which [illegible].

2) [illegible] only three did not [illegible] TH and TI is very high [illegible]).

3) Since the [illegible] TH without [illegible] corresponding [illegible] TI were [illegible] transfer-[illegible], the correlation [illegible] (.[illegible]) is larger than [illegible].

Since the [illegible] the criterion [illegible] is a purely [illegible] almost as good [illegible] using all possible variables.

In order [illegible] each variable was correlated with [illegible] all other [illegible] particled [illegible] partial correlations were [illegible] (.[illegible]), trainer [illegible] ([illegible]), trans-[illegible] TH (-.[illegible]), and TI (.[illegible]). [illegible] last [illegible] is significant [illegible] in the criterion, [illegible] correlation [illegible] correspondence between TH and [illegible] the criterion with all variables except [illegible] it [illegible] improbable that "[illegible] not [illegible] a common [illegible]" [illegible] over by a common phenomenon [illegible] this experimental situation, [illegible] a novel situation, [illegible] leads [illegible] which they have [illegible].

An analysis of variance (difficulty level by women [illegible] sex) of [illegible] the control group was also performed. The means, standard deviations and ranges for subjects are given in [illegible]. The summary of the results is given in Table [illegible]. Since there were no significant

Table 12

Means, Standard Deviations, Ranges, and Cell Frequencies of Trials to Criterion for Subgroups of the Control *Ss*

Group	Mean	n	S.D.	Range
"Easy" warmup task				
Male	17.4	8	4.6	12 - 24
Female	17.4	8	4.0	10 - 30
"Difficult" warmup task				
Male	19.8	8	4.0	13 - 23
Female	19.5	8	4.9	12 - 25

Table 13

Analysis of Variance of Trials to Criterion
for Subgroups of the Control Ss

Source	df	MS	F
Sex	1	0.3	0.01
Difficulty level (warmup task)	1	42.8	1.54
Difficulty level x Sex	1	0.3	0.01
Within	28	27.8	
Total	31		

Table 13

Analysis of Variance on Trials to Criterion for Subgroups of the Control Ss

Source	df	MS	F
Sex	1	0.8	0.01
Difficulty level (warmup task)	1	42.8	1.54
Difficulty level x Sex	1	0.3	0.01
Within	28	27.8	
Total	31		

differences, the data for the four groups were pooled in order to test controls against experimental Ss classified as not having consciously transferred. The latter Ss were those categorized as negative TI. Since there was not a significant difference between negative TI Ss trained with Rule 1 as opposed to Rule 2 (t = 1.59), all negative TI Ss were also pooled for the test against control Ss. There was no significant difference between control Ss and negative TI Ss (t = 0.79), and so no transfer without awareness appeared in the analysis of mean differences. However, four negative TI Ss solved the transfer task before the first task hint (given at trial 10) whereas none of the control Ss did so. Three of these negative TI Ss were trained with Rule 1, the other with Rule 2. All solved in 8 trials, except one who solved in 1 trial. He was trained with Rule 1. There was at least this one difference between the two groups.

Though there was no significant mean difference between control Ss given "easy" or "difficult" warmup items, there was a trend in the predicted direction with a mean of 17.4 for the "easy" group and 19.7 for the "difficult" group. Furthermore, 10 of the 16 Ss in the "easy" group solved the problem before the second task hint (given at trial 20), whereas only 4 of 16 Ss in the "difficult" group did so. Though this difference was not significant (x^2 = 3.18 $.10 > p > .05$), there is a suggestion of different difficulty sets, but it would have to be investigated further. This same trend appeared among negative TI Ss with a mean of 15.7 for those trained with Rule 1 and a mean of 19.3 for those trained with Rule 2. However, there might be another explanation for this difference. This explanation will be given later.

Type of solution. In E's task model, three rules solved the transfer task. All but two Ss in the experiment solved the task with one of these three rules. The two Ss who did not, however, used a variation of the 2TR [illegible], the data for the [illegible] groups were [illegible] to test control [illegible] that experimental [illegible] s not [illegible] consciously transformed [illegible] as [illegible] II. Since there [illegible] with Rule 1 [illegible] (t = 1.5[illegible]) [illegible] for the [illegible] control Ss. There was [illegible] between control Ss [illegible] (t = [illegible]), and [illegible] apparent [illegible] analysis [illegible] difference. However, [illegible] solve the transfer task before [illegible] whereas none of the control Ss did so. [illegible] with Rule 1, the other with Rule 2. [illegible] solved [illegible] trial. [illegible] difference between [illegible] groups.

Though there was no [illegible] difference between [illegible] given "easy" or "difficult" [illegible], there was a trend [illegible] direction with a [illegible] of 17. [illegible] "easy" group [illegible] "difficult" group [illegible] "easy" [illegible] between [illegible] the "difficult" group [illegible] (x^2 = 3.[illegible]) [illegible] sets, but it would [illegible] wrong [illegible] TI Ss [illegible] 13.7 [illegible] for [illegible] with [illegible]. However, [illegible] difference. This [illegible] will be [illegible] later.

[illegible] for the [illegible] solution. In [illegible] Ss solved the [illegible] task. All but two Ss in the experiment solved the [illegible] with [illegible] the three rules. The two Ss who did not, however, used a variation of the [illegible] rule. One used 2(T + R) - R; the other used T + R + T. Both were classified as solving with 2TR. Only three Ss found more than one solution; they found both of the 2 to 1 weighting solutions (2TR and 2CS). Their solutions were categorized according to the rule actually offered when forced to give one or the other.

Table 14 summarizes the frequency with which each type of solution was used by Ss in each of the experimental groups and in the control group. To test for differences in frequencies among experimental groups, the males and females in each group were pooled resulting in a highly significant difference ($X^2 = 55.72$, with 6 df, $p < .001$). Groups trained with Rule 1 and those trained with Rule 2 used different solutions; those trained with Rule 2 were the only Ss solving with 2CS.

The experimental groups were reclassified for further analysis. The no-transfer-hint groups were split into Ss with positive and negative TI. In Table 15, positive TI Ss from transfer-hint and no-transfer-hint groups are compared and negative TI Ss are compared with control Ss. Subjects trained with Rule 1 or Rule 2 were not pooled in these tables. The difference between positive TI Ss and either negative TI Ss or controls is obvious. Notice that the similarity of frequency split between positive TI Ss with or without a transfer hint and the same similarity between negative TI Ss and controls is always better for Ss trained with Rule 2. This difference between the two training rule groups will be discussed later.

Various Measures of the Learning Process

Various questions can be asked about each S's trial-by-trial hypotheses. This analysis looked for subtle differences in the learning process which analyses of trials to criterion and type of solution did not detect. For example, among Ss who spontaneously verbalized transfer information, the mean trial on which

Table 14

Frequency of Type of Solution in Each Experimental Group and Each Subgroup of Control Ss

Group	2TR	2CS	3CS
Rule 1			
Transfer hint			
Males	3	7	
Females	3	7	
Total	6	14	
No transfer hint			
Males	7	7	
Females	5	9	
Total	12	16	
Rule 2			
Transfer hint			
Males	3		7
Females	2	1	7
Total	5	1	14
No transfer hint			
Males	6	1	7
Females	10	1	3
Total	16	2	10
Control group			
"Easy" warmup			
Males	6	2	
Females	8		
Total	14	2	
"Difficult" warmup			
Males	6	2	
Females	8		
Total	14	2	

Table 15

Frequency of Type of Solution
for Positive TI (Transfer Hint vs. No Transfer Hint)
and Negative TI vs. Controls

Groups	Rules 2TR	2CS	3CS
Rule 1 - Positive TI			
Transfer hint	6	14	
No transfer hint	5	6	
Total	11	20	
Rule 2 - Positive TI			
Transfer hint	6		14
No transfer hint	3		10
Total	9		24
Negative TI			
Rule 1	11	6	
Rule 2	13	2	
Controls	28	4	

these verbalizations first occurred was 1.88 for the 33 Ss given the transfer hint, 3.83 for the 18 Ss not given the transfer hint. This mean difference suggested possible process differences, at least on the early trials. Furthermore, it was possible that the pretraining had effects among negative TI Ss, even though these effects were not sufficiently facilitating to produce a significant difference from the controls in the earlier analyses.

Information contained in the trial-by-trial hypotheses was divided into the following four categories:

1) The relationship of verbalized formulas to numerical answers. That is, was the agreement between the verbalized formulas and the numerical answers close enough so that the verbalized formulas could be analyzed as a reliable measure of the learning process?

2) Task model. How justified would E be in using his own task model to analyze performance? Are there detectable differences in the task model used by Ss in the various groups at different times in the learning process?

3) Known response scale. How much use did Ss make of the instruction about the response scale--that k could only be one of the whole numbers between 3 and 12?

4) Use of prior information. How much use of prior information, available through feedback, did Ss incorporate into their responses on a given trial?

Relationship of verbalized formulas to numerical answers. Including all experimental and control Ss, there were only 10 trials on which the verbalized formulas did not generate the numerical answer given. Considering that there was a total (across Ss) of 1537 trials before the criterion trials with formulas offered on 1278 of them, there was disagreement on only 00.8% of the trials. This was a remarkably high degree of agreement (99.2% of the trials), particularly since Ss were not allowed to use paper and pencil to calculate their responses and E did not point out miscalculations until after the response was given.

those wit[illegible] 33 Ss given the tran-fer hint [illegible] for hint. This mean dif-ference [illegible] test on the early trial [illegible]. Furthermore, it was possib[illegible] and effects among negat-[illegible], even though these [illegible]ly facilitating to produce a significant difference [illegible] in the earlier analyses.

Information contained in the [illegible] protocols was divided into the following four categories:

1) The relationship of verbalized [illegible] numerical answers. That is, was the agreement between [illegible] formulas and the numerical answers [illegible] verbalized formulas could [illegible] processing process?

2) Task model. [illegible] could E [illegible] task model to analyze [illegible] there [illegible] differences in the task model [illegible] Ss in the [illegible] different times in the learning process?

3) Known response scale. How much use [illegible] the instruction about the response scale--that R [illegible] of the whole numbers between 0 and 10?

4) Use of prior [illegible]. How [illegible] information, available through [illegible], did Ss [illegible] their responses on a given trial.

Relationship of verbalized [illegible] numerical answers. Including all experimental conditions, [illegible] trials on which the verbalized formulas did [illegible] were given. Considering [illegible] there was a total [illegible] the criterion [illegible] with formulas offered [illegible] disagreement on [illegible] of the trials. This was [illegible] agreement (90% of the trials), particularly since Ss were not [illegible] use paper to calculate their responses [illegible] pointed out miscalculations after the response was given.

Task model. Analysis of the task model was subdivided into analyses of 1) percent guessing (no formula given), 2) number of cues used, 3) type of cues used, 4) use of relative position, and 5) use of weighting.

Task model: percent guessing. The percent of Ss in each group who offered no formula was computed for each of the first two blocks of five trials. Percentages were based only on pre-criterion trials. For block 1 (trials 1 to 5) the percentages were 11% for positive TI (hint), 12% for positive TI (no hint), 11% for negative TI, and 24% for controls. For block 2 (trials 6 to 10) the percentages were 8% for positive TI (hint), 9% for positive TI (no hint), 15% for negative TI and 20% for controls. Since some of the control Ss had expressed confusion about what a "formula" might be, control and negative TI Ss in blocks 1 and 2 were compared. In block 1, the mean number of guesses for controls was 1.22, the mean for negative TI was 0.56 ($t = 2.18$, $p < .05$). In block 2, the mean for controls was 1.00, the mean for negative TI was 0.72 ($t = 0.90$). Control Ss guessed significantly more than negative TI Ss only in the first block of trials.

Task model: number of cues used. Number of cues refers to the number of figures (numerical value or position value) used in a formula. In the task model only two cues are relevant. Because the first task hint (at trial 10) specifically stated that only two cues are relevant, only the first 10 trials were analyzed. These 10 trials were divided into two blocks of 5 trials each. The dependent variable was the number of times only 2 cues were used. Means for each group in each block of 5 trials are given in Table 16. The analysis of variance (groups x blocks) is reported in Table 17. Group differences were significant beyond the .001 level ($F = 48.5$); there was no significant interaction. All groups were significantly different beyond the .01 level [illegible] model. Analysis of the [illegible] was subdivided into analyses of 1) percent guessing (no formula given) [illegible] number of guesses, [illegible] of cues used, 4) use of [illegible] [illegible] use of weighting.

Task model: percent guessing. [illegible] group who offered [illegible] the [illegible] trials. Percentages [illegible] (trials [illegible]) The [illegible] positive TI (no hint), [illegible] for negative [illegible] (trials [illegible]) the percentage [illegible] TI [illegible], [illegible] positive TI (no hint), 15% for negative [illegible] control [illegible] expected [illegible] control and [illegible] TI [illegible] [illegible] the mean number [illegible] [illegible] (t [illegible] p [illegible] .05). In [illegible] 2, [illegible] mean for control [illegible] the mean for [illegible] TI was [illegible] (t [illegible]). Control [illegible] significantly [illegible] than [illegible] TI [illegible] only [illegible].

Task model: number of cues used. [illegible] number of figures (numerical [illegible]) [illegible]. In the task model only two cues are relevant. [illegible] (at [illegible]) specifically state [illegible] first [illegible] analyzed. Thus [illegible] trials [illegible]. The dependent [illegible] means for each group in [illegible] Table 16. The analysis of variance (groups [illegible]) is reported [illegible] Table 17. Group differences [illegible] significant beyond the [illegible] level (F = 18.5); there was no significant interaction. All groups were significantly different beyond the .01 level

Table 16

Mean Number of Uses of Only Two Cues for Each Group in Each Block of Trials

Group	Block I (Trials 1 - 5)	Block II (Trials 6 - 10)	Total
Positive TI (hint)	3.75	4.32	8.07
Positive TI (no hint)	2.33	3.25	5.58
Negative TI	1.03	1.72	2.75
Controls	0.62	0.97	1.59

Table 17

Analysis of Variance of Number of Uses
of Only Two Cues

Source	df	SS	MS	F
Between SS	127	837.44		
Groups	3	452.27	150.76	48.5***
SS w. Groups	124	385.17	3.11	
Within SS	128	185.00		
Blocks	1	23.77	23.77	18.6***
Groups x Blocks	3	2.38	0.79	0.6
Blocks x SS w. Groups	124	158.85	1.28	
Total	255	1022.44		

***Significant beyond the .001 level

(Newman-Keuls procedure), with positive TI (hint) > positive TI (no hint) > negative TI > controls.

Task model: type of cues used. Fourteen different types of cues were used by Ss. These were:

1. Numbers in particular figures
2. Position values of particular figures
3. Numbers in figures at particular positions--regardless of the particular figure which appeared there
4. Figures with the same numbers inside them
5. Figures with different numbers inside them
6. The overall pattern of numbers
7. Arbitrary numbers assigned to figures
8. Highest or lowest numbers
9. Odd or even numbers
10. Number of sides of a particular figure
11. Number of corners of a particular figure
12. Figures whose inside number differed from its position value
13. Figures whose inside number was less than its position value
14. The card number (it was written on the back of the card)

These types of cues were used singly or in various combinations. Only the first two types belonged to the E's task model. When Ss used other than task model cues they used type-3 cues 86% of the time. The first 20 trials were divided into four blocks of five trials each. Means for the use of type-1, type-2 and other-than-task-model cues for all groups are presented in Table 18. Type-1 and type-2 cues were analyzed. In order to compensate for the fact that Ss who solved before trial 20 were assigned the type of cue in their solution rule for all subsequent trials, simple analyses of variance were performed for each block of trials. Therefore these were conservative analyses. The four analyses for type-1 cues are presented in Table 19. Groups were significantly different only in block 3 ($F = 6.24$, $p < .001$) and block 4 ($F = 6.21$, $p < .001$). The two positive TI groups were significantly different from the negative TI and control groups in both blocks 3 and 4 (Newman-Keuls procedure). There were no other significant differences.

Table 18

Mean Number of Uses of Type-1, Type-2, and Type-"Other" Cues for Each Group in Each Block of Trials

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Positive TI (hint)	1.34	1.20	1.24	1.29
Positive TI (no hint)	1.54	1.17	1.71	1.56
Negative TI	1.47	1.81	2.98	2.97
Controls	1.11	1.25	2.73	3.02

Type-2

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Positive TI (hint)	2.58	3.44	3.69	3.59
Positive TI (no hint)	1.77	2.52	3.08	3.35
Negative TI	0.73	0.53	1.08	1.09
Controls	0.20	0.34	0.38	0.31

Type-"other"

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Positive TI (hint)	0.76	0.24	0.02	0.10
Positive TI (no hint)	1.10	0.98	0.08	0.00
Negative TI	2.23	1.91	0.47	0.44
Controls	2.38	2.41	1.05	0.67

Table 19

Analyses of Variance of Type-1 Cues

in Each Block of Trials

Block 1

Source	df	SS	MS	F
Between	3	3.19	1.06	0.50
Within	124	260.49	2.10	
Total	127	263.68		

Block 2

Source	df	SS	MS	F
Between	3	8.86	2.95	1.08
Within	124	339.61	2.74	
Total	127	348.47		

Block 3

Source	df	SS	MS	F
Between	3	71.18	23.73	6.24***
Within	124	471.68	3.80	
Total	127	542.86		

Block 4

Source	df	SS	MS	F
Between	3	83.25	27.75	6.21***
Within	124	554.81	4.47	
Total	127	638.06		

***Significant beyond the .001 level

The four analyses for type-2 cues are presented in Table 20. Groups were significantly different in block 1 (F = 16.31, $p < .001$), block 2 (F = 31.36, $p < .001$), block 3 (F = 25.83, $p < .001$) and block 4 (F = 24.50, $p < .001$). Using the Newman-Keuls procedure, the two positive TI groups were significantly different from the negative TI and control groups in all four blocks of trials. Positive TI (hint) was significantly different from positive TI (no hint) beyond the .05 level in both blocks 1 and 2. There were no other significant differences.

The four analyses for type-"other" are presented in Table 21. Groups were significantly different in block 1 (F = 14.88, $p < .001$), block 2 (F = 19.3, $p < .001$), block 3 (F = 9.99, $p < .001$) and block 4 (F = 5.43, $p < .01$). Using the Newman-Keuls procedure, the positive TI groups were significantly lower than negative TI and controls in blocks 1 and 2. In block 2, positive TI (hint) was significantly lower than positive TI (no hint). In block 3, all groups were significantly lower than the controls. In block 4, positive TI (no hint) was significantly lower than both negative TI and the controls; positive TI (hint) was significantly lower than the controls.

Though the negative TI and control groups did not differ significantly in use of type-1 or type-2 cues in any block of trials, the negative TI had a higher mean in all blocks of trials for both types of cues except for type-1 in block 4. Therefore these two types of cues were combined for these two groups in all four blocks and one-tailed t-tests were used to assess differences. The means for negative TI Ss in the four blocks were 2.20, 2.34, 4.06 and 4.06; the means for controls were 1.31, 1.59, 3.11 and 3.33. All means were significantly different. In block 1, t = 2.24 ($p < .05$), in block 2, t = 1.75 ($p < .05$), in block 3, t = 2.41 ($p < .01$), and in block 4, t = 1.81

Table 20

Analyses of Variance of Type-2 Cues

in Each Block of Trials

Source	df	SS	MS	F
Between	3	118.43	39.48	16.31***
Within	124	300.69	2.42	
Total	127	419.12		

Block 2

Source	df	SS	MS	F
Between	3	238.95	79.65	31.36***
Within	124	314.52	2.54	
Total	127	553.47		

Block 3

Source	df	SS	MS	F
Between	3	252.66	84.22	25.83***
Within	124	403.72	3.26	
Total	127	656.38		

Block 4

Source	df	SS	MS	F
Between	3	260.90	86.97	24.50***
Within	124	439.78	3.55	
Total	127	700.68		

***Significant beyond the .001 level

Table 21

Analyses of Variance of Type-"Other" Cues in Each Block of Trials

Block 1

Source	df	SS	MS	F
Between	3	66.06	22.02	14.88***
Within	124	183.97	1.48	
Total	127	250.03		

Block 2

Source	df	SS	MS	F
Between	3	98.38	32.79	19.3***
Within	124	210.67	1.70	
Total	127	309.05		

Block 3

Source	df	SS	MS	F
Between	3	21.57	7.19	9.99***
Within	124	89.46	0.72	
Total	127	111.03		

Block 4

Source	df	SS	MS	F
Between	3	8.78	2.93	5.43**
Within	124	67.28	0.54	
Total	127	76.06		

**Significant beyond the .01 level

***Significant beyond the .001 level

($p < .05$). Negative TI Ss consistently used task-model cues more than the control Ss.

There was one further point of interest about the use of cues. Among negative TI Ss and controls, the use of non-task-model cues (type-"other") dropped considerably after trial 10. Since the first task hint was given at this time, presumably that hint gave more information to these Ss than the fact that only two cues are relevant. Apparently, at least to these Ss, the hint gave information about both the number and the type of cues.

Task model: use of relative position. Relative position is defined as the incorporation, in some way or other, of the position of one figure in respect to another (or others) into a formula. Though relative position in the task model is related to adding or subtracting a constant, some of the actual uses of relative position were related to different formulas, applied conditionally depending upon which of two figures was "to the left." Since relative position was used only four times by controls, once by negative TI Ss, and twice by positive TI Ss trained with Rule 1, only the positive TI Ss trained with Rule 2 in the hint and no-hint groups were compared. The first 20 trials were again divided into four blocks of five trials each. Means for each group in each block are given in Table 22. A groups x blocks analysis of variance is also reported in Table 22. Neither groups nor groups x blocks was significant.

Task model: use of weighting. Weighting was analyzed in terms of two categories: 1) any type of weighting, other than 1, 2) task-model weighting. Means for all groups for each category are presented in Table 23. There means represent the mean trial on which this type of weighting was first used. Analyses of variance are also presented in Table 23. The groups were significantly different both for any type of weighting ($F = 28.97$, $p < .001$) and for task-

(p < .05). [illegible] consistently [illegible] task-model cues more than the control Ss.

There [illegible] about the use of cues. [illegible] negative TI [illegible] task-model cues (type-" [illegible]) dropped considerably after [illegible] first task hint was given at this time, presumably [illegible] information to these Ss that the [illegible] apparently, at least to these Ss, [illegible] the [illegible] cues.

Task [illegible] use of [illegible] positive [illegible] the [illegible] of [illegible] respect [illegible] relative [illegible] the [illegible] constant, [illegible] formula, [illegible] Since relative [illegible] TI [illegible], and [illegible] positive TI [illegible] defined [illegible] first 24 trials were again divided [illegible] for each group in each block are [illegible] analysis of variance is [illegible] reported in [illegible] blocks was significant.

Task [illegible] analyzed in terms of [illegible] gories: 1) [illegible] 2) task-model weighting. [illegible] for all groups [illegible] in Table 23. [illegible] the [illegible] on which this type of weighting was first used. Analyses of [illegible] also presented in Table 23. The groups were significantly different both for any type of weighting (F = 25.97, p < .001) and for task-

Table 22

Means for the Use of Relative Position in Each Block of Trials for Positive TI (Hint) and Positive TI (No Hint) Groups Trained with Rule 2

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Positive TI (hint)	2.70	3.65	3.90	3.70
Positive TI (no hint)	2.54	3.23	3.62	3.92

Analysis of Variance of the Use of Relative Position by Positive TI (Hint) and Positive TI (No Hint) Groups Trained with Rule 2

Source	df	SS	MS	F
Between SS	32	414.74		
Groups	1	0.81	0.81	0.06
SS w. Groups	31	413.93	13.35	
Within SS	99	155.50		
Trials	3	29.33	9.78	7.30***
Groups x Trials	3	1.81	0.60	0.45
Trials x SS w. Groups	93	124.36	1.34	
Total	131			

***Significant beyond the .001 level

Table 23

Mean Trials on Which Any Type of Weighting and Task Model Weighting was Used by All Groups

Groups	Any type of weighting	Task-model weighting
Positive TI (hint)	2.69	3.60
Positive TI (no hint)	5.17	7.08
Negative TI	11.31	14.66
Controls	12.44	15.75

Analyses of Variance of Trials on Which

Any Type of Weighting and Task-Model Weighting

were First Used by All Groups

Any Type of Weighting

Source	df	SS	MS	F
Between	3	2299.63	766.54	28.97***
Within	124	3280.86	26.46	
Total	127	5580.49		

Task-Model Weighting

Source	df	SS	MS	F
Between	3	3593.97	1197.99	41.79***
Within	124	3554.65	28.67	
Total	127	7148.62		

***Significant beyond the .001 level

model weighting (F = 41.79, p < .001). Using the Newman-Keuls procedure, both positive TI groups used both types of weighting significantly sooner than the negative TI and control groups. Positive TI (hint) used task-model weighting significantly sooner than positive TI (no hint).

Known response scale. Since *S*s were told the *k*-scale, that correct answers could only be some whole number between 3 and 12, they could use that information in two ways. First, they could eliminate any response of numbers less than 3 or greater than 12. Second, they could monitor any formula by checking to see whether it generated all and only whole numbers between 3 and 12.

Subjects gave responses that ranged from -1 to +40. The number of off-scale responses were: positive TI (hint) 6, positive TI (no hint) 8, negative TI 25, and controls 10. Thirty-two of these off-scale responses were numbers less than three; 17 were numbers greater than 12. Since *E* reminded *S* of the *k*-scale when an off-scale response was given, these frequencies were probably somewhat controlled.

When the possible answers a formula could generate were compared with the *k*-scale, the verbalized formulas could be categorized into five groups. Because there was little difference among the four basic groups of *S*s, they were pooled. The five categories of formulas and the percents in each of them were:

1) Give off-scale answers less than 3 (26%)

2) Give off-scale answers greater than 12 (13%)

3) Give off-scale answers both less than 3 and greater than 12 (25%)

4) Give only on-scale answers, but not all of them (7%)

5) Give all and only on-scale answers (29%)

Since only 29% of the formulas used seemed to take the known *k*-scale into model weighting (F = 11.67, p < .001). Using the Newman-Keuls procedure, both positive TI groups used both types of weighting significantly sooner than the negative TI and control groups. Positive TI (hint) used task-model weighting significantly sooner than positive TI (no hint).

Known response limits. Since Ss were told the K-scale, that correct answers could only be a whole number between 3 and 12, they could use that information in two ways. First, they could eliminate any response that was less than 3 or greater than 12. Second, they could monitor a formula by checking to see whether it generated all and only whole numbers between 3 and 12.

Subjects gave responses that ranged from -1 to +40. The numbers of off-scale responses were: positive TI (hint) 6, positive TI (no hint) 8, negative TI 25, and controls 10. Thirty-two of those off-scale responses were numbers less than three; 17 were numbers greater than 12. Since E reinstated L of the K-scale when an off-scale response was given, these frequencies were probably somewhat controlled.

When the possible answers a formula could generate were compared with the K-scale, the verbalized formulas could be categorized into five groups. Because there was little difference among the four basic groups of Ss, they were pooled. The five categories of formulas and the percents in each of them were:

1) Give off-scale answers less than 3 ([illegible])

2) Give off-scale answers greater than 12 ([illegible])

3) Give off-scale answers both less than 3 and greater than 12 ([illegible])

4) Give only on-scale answers, but not all of them ([illegible])

5) Give all and only on-scale answers ([illegible])

Since only [illegible] of the formulas used seemed to take the known K-scale into account, _S_s apparently made little use of this possible information. Only a few _S_s spontaneously verbalized the use of such information.

Use of prior information. Three different questions can be asked about each formula verbalized by an _S_. Has this formula already been disproved, either explicitly or implicitly? Would this formula have given the correct answer on the preceding trial? What was _S_'s subjective estimate of whether his formula would have worked on the preceding trial? In interpreting this data, it should be remembered that the task was complex and _S_'s memory load was at a maximum since he saw only one card at a time and could not record prior information on scrap paper.

Use of prior information: status of the formula. Incorrect formulas were divided into the following three categories:

1) Insufficient information to reject it--though false, it would have given the correct answer on all previous cards.

2) Sufficient information but _implicit_--the formula would not have given the correct answer on one or more previous cards, but _S_ had not explicitly used it.

3) Sufficient information and _explicit_--the formula was previously used by the same _S_ and had not given the correct answer.

Since the differences between groups were slight, overall percentages for the three categories were computed. There were 16% of the formulas in category 1, 79% in category 2, and 5% in category 3. For category 2, the mean number-of-trials-back that the formula was implicitly disconfirmed was 1.84. This means that, on the average, the _S_s were taking account of approximately the last two cards. For category 3, the mean number-of-trials-back that the formula was explicitly disconfirmed was 4.6. The fact that category 2 formulas were disconfirmed by a card which appeared less than 2 cards earlier on the average and the fact that 79% of the incorrect formulas were in this category suggests a severe memory problem.

account, Ss apparently [illegible] little use of this possible [illegible]ation. Only a few Ss spontaneously verbalized th[illegible] such information.

Use of prior information. Thr[illegible] questions can be asked about each formula verbalized by an S. [illegible] ready been dis[illegible] either explicitly or implicitly [illegible] this [illegible] have given the correct answer on the preceding trial? What was the [illegible] estimate of whether his formula would have [illegible] on the preceding [illegible] In interpreting this data, it should be [illegible] that the task [illegible] and S's memory load was at a maximum since [illegible] only one card [illegible] and could not record prior information on scrap paper.

Use of prior information: status of the formula. Incorrect formulas were divided into the following [illegible] categories:

1) Insufficient information [illegible] it--though [illegible], it would have given the [illegible] all previous cards.

2) Sufficient information [illegible] it--the formula would not have given the correct answer on one or more previous cards, but S had not explicitly [illegible]

3) Sufficient information [illegible] it--the formula was previously used by the same S [illegible] for the correct answer.

Since the differences between [illegible] were slight, overall percentages for the three categories were [illegible]. There were [illegible] the formulas in category 1, 79% in category [illegible] in category [illegible] For category 2, the mean number-of-trials-back that the formula was implicitly [illegible] discounted was 1,84. This means that, on the average, the Ss were taking account of approximately the last two cards. For category 3, the mean number-of-trials-back that the formula was explicitly [illegible] was 1.6. The fact that category [illegible] formulas were disconfirmed by [illegible] less than [illegible] earlier on the average and the fact that [illegible] the incorrect formulas were in this category suggests [illegible] seve[illegible] memory [illegible].

Use of prior information: formula on trial n in respect to the formula on trial n-1. How much did Ss use the information from the immediately preceding trial? The percentages of formulas that would have worked or would not have worked on the previous card are given in Table 24. Pooling all groups, Ss did not make use of the information on the previous trial on 29% of the trials. The largest difference was between positive TI (hint) with 35% and positive TI (no hint) with 24%. It is possible that the high percentage for positive TI (hint) was related to the transfer hint. Possibly these Ss were paying some attention to using a formula related to the training task, and therefore payed less attention to the feedback for the preceding trial. There seems to be no obvious reason why the percentage of positive TI (no hint) should be the lowest.

Use of prior information: S's subjective estimate of whether the formula on trial n worked on trial n-1. For most trials information was available as to whether S thought he was using a formula which would have worked on the preceding trial. These subjective estimates can be categorized into six rough categories, no matter whether the formula actually did work on the preceding card or not. The six categories are:

a) S was sure it worked.
b) S "thought" it worked.
c) S was not sure or didn't remember.
d) S was sure it did not work.
e) No information given.
f) S claimed to have guessed, though he offered a formula.

The information about these categories is also given in Table 24. Subjects seemed to be more sure of themselves if the formula had worked on the previous card than if it had not worked. And though no S ever said a formula had not worked on the previous card when it actually had, Ss in all groups occasionally thought the formula they were using had worked on the previous card when it actually had not.

Use of prior information: [illegible] n in respect to the formula on trial n-1. How much did Ss [illegible] information from the immediately preceding trial? The percentages [illegible] have worked or would not have worked on the previous [illegible] Table 24. Pooling all groups, Ss did not [illegible] the previous trial in 40% of the trials. The largest [illegible] positive TI [illegible] with RRI and positive TI [illegible]. It is [illegible] that the [illegible] percentage for positive TI (high [illegible]) [illegible]. Possibly these Ss were paying some attention [illegible] the training task, and therefore [illegible] the feedback for the preceding trial. [illegible] positive TI [illegible] the lowest.

[illegible] estimate whether the formula [illegible] worked on [illegible] trial [illegible] information was available as to [illegible] which would have worked on the [illegible]. The [illegible] estimates can be categorized into six rough categories, [illegible] formula actually did work on the preceding [illegible] S [illegible] are:

1) S was sure [illegible].
2) S "thought" it worked.
3) S was not sure [illegible].
4) [illegible] sure [illegible].
5) S was [illegible] given.
6) No guess [illegible] formula.

The information about [illegible] given in Table 25. Subjects seemed to be more sure [illegible] worked on the previous card than if it had not worked. [illegible] no S [illegible] worked on the previous card [illegible] had, [illegible] a formula had not [illegible] occasionally thought the [illegible] they were using had worked on the previous card when it actually had not.

Table 24

Percent of Subjective Estimates in Six Categories

for Formulas which Would or Would Not have Worked on the Preceding Trial

for Each of the Four Basic Groups

Groups	Formulas which would have worked						Formulas which would not have worked					
	a	b	c	d	e	f	a	b	c	d	e	f
Positive TI (hint)	64	1	0	0	0	0	5	1	15	6	7	1
Total percent				65%						35%		
Positive TI (no hint)	72	0	0	0	2	2	8	0	5	5	3	3
Total percent				76%						24%		
Negative TI	69	1	1	0	0	0	5	2	8	6	2	5
Total percent				72%						28%		
Controls	68	1	1	0	0	0	5	1	7	9	4	3
Total percent				71%						29%		
Overall percent				71%						29%		

Effects of "easy" and "difficult" pretraining. Controls Ss were trained with "easy" or "difficult" items of Raven's Progressive Matrices Test since there was some reason to believe that a differential set for difficulty level might have been developed in negative TI Ss trained with Rule 1 ("easy" set) or with Rule 2 ("difficult" set). These subgroups were compared on various measures of the learning process. Since controlling for different difficulty-level sets was only a hypothesis, fairly conservative tests for differences were used.

Amount of guessing. The mean number of guesses for each group in each of the first two blocks of trials is given in Table 25. A simple analysis of variance for each block of trials showed no significant differences. For block 1, $F = 2.49$, $p < .05$; for block 2, $F = 0.33$.

Number of cues used. The mean number of use of two-cues for each group in each of the first two blocks of trials is given in Table 26. There was no significant difference between groups in block 1 ($F = 0.96$). There was a significant difference between groups in block 2 ($F = 3.41$, $p < .01$). Negative TI Ss trained with the "easy" rule (Rule 1) used two cues in block 2 significantly more than both the negative TI Ss trained with the "difficult" rule (Rule 2) and the control Ss with "difficult" pretraining (Newman-Keuls procedure).

Type of cues used. The mean number of times type-1, type-2, and type-other-than-task-model cues was used by each group in each block of trials is given in Table 27. There was no difference in the use of type-1 cues in any block of trials. The F's for the four blocks were 2.09, 0.93, 1.27 and 0.11 respectively ($p > .05$ for all).

The Fs for the use of type-2 cues in each block of trials was 3.67, 1.76, 3.74 and 3.02. The Fs for blocks 1, 3 and 4 were significant beyond the .05

Table 25

Mean Number of Guesses in Each Block of Trials for the "Easy" and "Difficult" Pretraining Subgroups of Negative TI and Control *Ss*

Groups	Block 1	Block 2
Easy (controls)	1.56	0.94
Difficult (controls)	0.88	1.06
Easy (Rule 1)	0.65	0.65
Difficult (Rule 2)	0.47	0.80

Table 26

Mean Number of Uses of Two-Cues in Each Block of Trials for the "Easy" and "Difficult" Pretraining Subgroups of Negative TI and Control Ss

Groups	Block 1	Block 2
Easy (controls)	0.56	1.19
Difficult (controls)	0.69	0.75
Easy (Rule 1)	1.00	2.18
Difficult (Rule 2)	1.07	1.20

Table 27

Mean Number of Uses of Type-1, Type-2, and Type-"Other" Cues in Each Block of Trials by "Easy" and "Difficult" Control or Negative TI Groups

Type-1

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Easy (control)	0.91	1.16	3.00	3.22
Difficult (control)	1.31	1.34	2.47	2.81
Easy (Rule 1)	0.97	1.59	2.50	2.94
Difficult (Rule 2)	2.03	2.07	3.53	3.00

Type-2

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Easy (control)	0.25	0.47	0.28	0.25
Difficult (control)	0.16	0.22	0.47	0.38
Easy (Rule 1)	0.97	0.85	1.62	1.53
Difficult (Rule 2)	0.47	0.17	0.47	0.60

Type-"other"

Groups	Block I Trials 1-5	Block II Trials 6-10	Block III Trials 11-15	Block IV Trials 16-20
Easy (control)	2.22	2.44	0.59	0.41
Difficult (control)	2.53	2.38	1.50	0.94
Easy (Rule 1)	2.41	1.91	0.35	0.35
Difficult (Rule 2)	2.03	1.90	0.60	0.53

level. Negative TI Ss trained with the "easy" rule used type-2 cues significantly more than control Ss in the "easy" group in blocks 1, 3 and 4, than control Ss in the "difficult" group in blocks 1 and 3, than negative TI Ss trained with the "difficult" rule in block 3 (Newman-Keuls procedure). There were no other significant differences.

The Fs for the use of other-than-task-model cues in each block of trials were 0.48, 0.60, 3.21 and 1.23 respectively. Only the F for block 3 was significant beyond the .05 level. In that block, control Ss with "difficult" pretraining used this type of cue significantly more than the negative TI Ss trained with either the "easy" or "difficult" rule (Newman-Keuls procedure). There were no other significant differences.

Use of weighting. The mean trials on which each group first used any type of weighting (other than 1) and task-model weighting is given in Table 28. There was no significant difference between groups for "any" type of weighting ($F = 1.08$). There was also no significant difference between groups for the first use of task-model weighting ($F = 1.56$).

Discussion and Conclusions

The results of this experiment will be discussed in terms of the following categories: 1) the effect of transfer intentions (positive TI Ss compared with either negative TI Ss or controls), 2) the effect of a transfer hint (including a comparison of positive TI Ss either given this hint or not), 3) transfer-without-awareness effects (a comparison of negative TI Ss with controls), and 4) set-for-difficulty-level effects (a comparison of negative TI Ss and controls who had "easy" or "difficult" pretraining).

level. [illegible] used with the "easy" rule used two-3 cues signifi-
cantly [illegible] $\underline{S}$s in the "easy" group in blocks 1, [illegible] and 4, than
[illegible] "difficult" group [illegible] 1 and 3, than negative TI $\underline{S}$s
trained with the "difficult" [illegible] (Newman-Keuls procedure). There
were [illegible] significant [illegible].

[illegible] $\underline{S}$s [illegible] other-than-task-[illegible] cues in each block [illegible]
[illegible], 8.12 and [illegible] respectively. Only the $\underline{F}$ for block 3 was
significant beyond the [illegible] level. In this block, [illegible] $\underline{S}$s with "difficult"
pretraining [illegible] the negative TI
$\underline{S}$s trained with [illegible] "easy" [illegible] "difficult" [illegible] (Newman-Keuls procedure).
There were no other significant [illegible].

[illegible] weighting. The [illegible] each group first used any
type of weighting (other than [illegible]) [illegible] weighting is given in Table
28. There [illegible] significant [illegible] between groups for "any" type of
weighting ($\underline{F}$ = [illegible]). There was [illegible] significant difference between groups
for the first use of [illegible] weighting ($\underline{F}$ = [illegible]).

Discussion and Conclusions

[illegible] results of this experiment will be discussed [illegible] terms of the follow-
ing [illegible]: 1) the effect of transfer intentions ([illegible]tive TI $\underline{S}$s compared
with either negative TI [illegible] controls), 2) the effect of a transfer hint
(including a comparison [illegible] TI $\underline{S}$s either given [illegible] hint or not),
3) transfer [illegible] effects [illegible] TI $\underline{S}$s with
controls), and 4) [illegible] comparison of negative
TI $\underline{S}$s and controls who had "easy" or "difficult" [illegible]).

Table 28

Mean Trial on Which Any Type of Weighting (Other than 1) and Task-Model Weighting were First Used by "Easy" and "Difficult" Control Groups or Negative TI Group

Groups	Any type of weighting	Task-model of weighting
Easy (control)	12.06	15.50
Difficult (control)	12.81	16.00
Easy (Rule 1)	9.65	12.82
Difficult (Rule 2)	13.20	16.73

The Effect of Transfer Intentions

Trials to criterion. Mean trials to criterion for positive TI Ss, negative TI Ss and controls were 6.3, 17.4, and 18.5 respectively. The superiority of positive TI Ss is obvious. Furthermore, among experimental Ss, only transfer intention was intrinsically related to trials to criterion when all other variables were partialled out.

Type of solution. The percentage of solutions of the transfer task by means of rules identical to those given in the training task was much higher for positive TI Ss. The only Ss who solved the transfer task with the 3CS rule were positive TI Ss who had been trained with Rule 2.

Learning process. a) Positive TI Ss used two cues more than negative TI Ss and controls in the first two blocks of trials. b) Positive TI Ss used type-1 cues less in blocks 3 and 4, and type-2 cues more in all four blocks of trials. They used type-"other" cues less in blocks 1 and 2 than both negative TI Ss and controls. In block 3, both positive TI Ss given the transfer hint and those who were not used type-"other" cues less than controls. In block 4, positive TI (hint) used type-"other" less than controls whereas positive TI (no hint) used them less than both negative TI Ss and controls. Since type-2 cues were similar to those used in the pretraining task, their more frequent use by positive TI Ss is related to the fact that they were transferring. c) Only positive TI Ss trained with Rule 2 used "relative position" in any substantial amount. Only this group should have used "relative position" frequently since they were consciously transferring and had been trained with a rule which included the use of "relative position." d) Positive TI Ss used "any" type of weighting and task-model weighting sooner than negative TI Ss and controls. This earlier use of weighting, especially

task-model weighting, reflects the fact that these Ss were transferring.

The Effect of the Transfer Hint

There was a significant transfer-hint effect in trials to criterion in the analysis of the basic experimental groups. However transfer hint was not intrinsically related to trials to criterion when the cognitive transfer processes (TH and TI) were partialled out. The locus of effect of the transfer hint was in the occurrence of positive transfer hypotheses and transfer intentions in all members of the transfer-hint group. There were significantly more positive TH and positive TI Ss in this group than in the experimental group not given this transfer hint, and this difference accounted for the difference in trials to criterion for the two groups.

There were no significant differences between positive TI Ss given a transfer hint and those who were not in either trials to criterion or type of solution. The following differences were found in the learning process measures: a) Transfer-hint Ss used two cues more in the first two blocks of trials. b) Transfer-hint Ss used type-2 cues more in blocks 1 and 2, and type-"other" less in block 2. c) Transfer-hint Ss used task-model weighting sooner than positive TI Ss not given the transfer hint. These three differences suggest that the transfer hint caused positive transfer hypotheses and intentions to occur sooner. The fact that, among Ss who spontaneously verbalized transfer information, these verbalizations first occurred on a mean trial of 1.88 for transfer-hint Ss and a mean trial of 3.83 for no-transfer-hint Ss suggests this same conclusion. An earlier occurrence of transfer hypotheses and intentions for positive TI Ss given the transfer hint was not predicted. Its effect was not enough to task-model weighting, reflects the fact that these Ss were transferring.

<u>The Effect of the Transfer Hint</u>

There was a significant transfer-hint effect in trials to criterion in the analysis of the basic experimental group. However transfer hint was not statistically related to trials to criterion when the cognitive transfer processes (TH and TI) were partialled out. The locus of effect of the transfer hint was in the occurrence of positive transfer hypotheses and transfer intentions in all members of the transfer-hint group. There were significantly more positive TH and positive TI Ss in this group than in the experimental group not given this transfer hint, and this difference accounted for the difference in trials to criterion for the two groups.

There were no significant differences between positive TI Ss given a transfer hint and those who were not in either trials to criterion or type of solution. The following differences were found in the learning process measures: a) transfer-hint Ss used two cues more in the first two blocks of trials, b) transfer-hint Ss used type-3 cues more in blocks 1 and 2, and type-"other" less in block 2. c) Transfer-hint Ss used task-model weighting sooner than positive TI Ss not given the transfer hint. These three differences suggest that the transfer hint caused positive transfer hypotheses and intentions to occur sooner. The fact that, among Ss who spontaneously verbalized transfer intentions, these verbalizations first occurred on a mean trial of 1.88 for transfer-hint Ss and a mean trial of 3.83 for non-transfer-hint Ss suggests this same conclusion. An earlier occurrence of transfer hypotheses and intentions for positive TI Ss given the transfer hint was not predicted. Its effect was not enough to show up in a significant difference in trials to criterion.

One further small difference was detected. The three Ss who reported positive TH but negative TI were all in the no-transfer-hint group. If this effect is reliable, it might suggest that transfer-hint Ss have more certainty while pursuing a transfer hypothesis than no-transfer-hint Ss.

Transfer-Without-Awareness Effects

There was not a significant difference between negative TI Ss and controls in either trials to criterion or type of solution. However, four negative TI Ss solved the transfer task before the first task-hint, whereas none of the control Ss did so. The following differences in measures of the learning process were found: a) The controls did more guessing than negative TI Ss in block 1. This difference seems related to the puzzlement of some control Ss about "what a formula might be." Controls had not had experience with the use of a formula in pretraining, whereas negative TI Ss had such experience. b) Negative TI Ss used two cues more in the first two blocks of trials. c) Negative TI Ss used type-"other" cues less than controls in block 3. And when type-1 and type-2 cues were pooled, negative TI Ss used them significantly more than controls in all blocks of trials.

The difference in (a) seems to be a non-specific type of transfer if specific is related to the use of some element of the pretraining rule. The differences in (b) and (c) suggest that negative TI Ss, at least in respect to the number and type of cues used, found E's task model sooner. These differences also seem to be a type of non-specific transfer which is not accounted for by conscious relating of the two tasks. The use of only two sho[illegible] errors in trials to criterion)

[illegible] difference was detected. The third Ss who reported positive [illegible] TI more [illegible] in the non-transferring group. If this [illegible], it might suggest that transferring [illegible] more [illegible] transfer [illegible] than non-transferring Ss.

Transfer-[illegible] Extraneous Effects

There was no significant difference between negative TI Ss and controls in [illegible] trials to criterion or type of solution. However, four negative [illegible] transfer [illegible] on the first task, whereas none of the controls did so. The following differences in measures of the learning process [illegible]: a) The controls did more cues [illegible] using the negative TI [illegible] in block 1. This difference seems related to the [illegible] ment of negative Ss about "what a [illegible] might be." Controls had not had experience with the use of a rule in pretraining, whereas negative TI Ss had such experience. b) Negative TI Ss used two cues more in the first two blocks of trials. c) Negative TI Ss used type-"other" cues less than controls in block 2. And when type-1 and type-2 cues were pooled, negative TI Ss used them significantly more than controls in all blocks of trials.

The difference in (a) seems to be a non-specific type of transfer in specific is related to the use of some element of the pretraining rule. The differences in (b) and (c) suggest that negative TI Ss, at least in respect to the number and type of cues used, found E's task model sooner. These differences also seem to be a type of non-specific transfer which is not accounted for by considering [illegible] of the two tasks. The use of only two cues demonstrates a willingness to ignore information, and perhaps experience with the training task in which such ignoring was done facilitated this type of behavior in the transfer task. These differences between negative TI Ss and controls were not predicted.

Set for Difficulty Level

Control Ss were divided into "easy" or "difficult" pretraining by giving either the easy or difficult items of Raven's Progressive Matrices Test. Since experimental Ss trained with Rule 2 took longer to read the training instructions, took longer to attain criterion in the training task, and made more errors in the training task, Rule 2 was more difficult to learn and use. Therefore, it was possible that different sets for difficulty level developed in each subgroup of control and negative TI Ss.

There were no significant differences in mean trials to criterion between the four subgroups, nor was there any difference in type of solution. However, trials to criterion were in the predicted direction and when the number of Ss in each subgroup who solved before the second-task hint were compared, negative TI Ss trained with Rule 1 (easy) had significantly more, and the "easy" controls almost had significantly more, than the "difficult" controls. There were also no differences in guessing between the four subgroups, even though the pooled controls guessed more in block 1 than the pooled negative TI Ss.

The following differences between these subgroups in measures of the learning process were found: a) Negative TI Ss trained with Rule 1 (easy) used two cues in block 2 more than negative TI Ss trained with Rule 2 (difficult) and the "difficult" controls. b) Negative TI Ss trained with Rule 1 (easy) used type-2 cues more than either control subgroup in block 1,

...ous demonstrated a willingness to ignore information, and perhaps experience with the training task in which such ignoring was done facilitated this type of behavior in the transfer task. These differences between negative TI Ss and controls were not predicted.

Set for Difficulty Level

Control Ss were divided into "easy" or "difficult" pretraining by giving either the easy or difficult items of Raven's Progressive Matrices Test. Since experimental Ss trained with Rule 2 took longer to read the training instructions, took longer to attain criterion in the training task, and made more errors in the training task, Rule 2 was more difficult to learn to use. Therefore, it was possible that different sets for difficulty level developed in each subgroup of control and negative TI Ss.

There were no significant differences in mean trials to criterion between the two subgroups, nor was there any difference in type of solution. However, trials to criterion were in the predicted direction and when the number of Ss in each subgroup who solved before the second-task hint were compared, negative TI Ss trained with Rule 1 (easy) had significantly more, and the "easy" control subset had significantly more, than the "difficult" controls. There were also no differences in guessing between the two subgroups, even though the pooled controls guessed more in block 1 than the pooled negative TI Ss.

The following differences between these subgroups in measures of the learning process were found: a) Negative TI Ss trained with Rule 1 (easy) used [illegible] cues in block 2 more than negative TI Ss trained with Rule 2 (difficult) and the "difficult" controls. b) Negative TI Ss trained with Rule 1 (easy) used type-2 cues more than either control subgroup in block 1.

more than any other subgroup in block 3, and more than the "easy" controls in block 4. The "difficult" control group used type-"other" cues more in block 3 than either of the negative TI subgroups.

Therefore, there were some differences that could be attributable to difficulty level. Negative TI Ss trained with the "easy" role used two cues more than the other groups, significantly more than the "difficult" groups. This might be attributable to a set for an easier problem since they were willing to use less information. However, their greater use of type-2 cues might also suggest some type of transfer-without-awareness effect, which could explain that group's superiority. A higher percentage of Ss in this particular subgroup did solve the transfer task with the 2CS rule. The fact that the "difficult" control group performed worst in terms of solving before the second task hint and the measures of the learning process seems to have no other explanation that difficulty set.

When trials to criterion alone was used as the criterion of transfer, it appeared that there were no transfer-without-awareness effects and that transfer of a specific rule could be explained solely in terms of positive transfer hypotheses and transfer intentions. Using type of solution as a further criterion of transfer did not change this general conclusion. However, the various measures of the learning process suggested some differences between positive TI Ss who either were or were not given a transfer hint, some transfer-without-awareness effects, and some set-for-difficulty-level effects. Thus, measures of the learning process seem to be the most powerful criterion for detecting transfer effects. These differences detected by measures of the learning process in no way detract from the powerful effect of the cognitive transfer processes.

more than any other subgroup in [illegible] 3, not more than the 'easy' controls in [illegible]. The "[illegible]" control group used type-II [illegible] more in [illegible] TI [illegible].

Therefore [illegible] alternatives that [illegible] attributable to [illegible] TI is trained with the "[illegible]" two ages more than [illegible] other [illegible] significantly more than the [illegible] used [illegible] groups. [illegible] problem since [illegible] [illegible]. However, their greater use of type-I rules [illegible] of transfer-rather-awareness effect, which could explain [illegible] A higher percentage of Ss in this particular [illegible] the transfer task with the 2CS rule. The fact that the [illegible] worst in terms of solving [illegible] learning process seem [illegible] other explanation [illegible].

[illegible] as the criterion of transfer, it appears that there were [illegible] effects and that transfer of [illegible] in terms of positive transfer hypotheses and transfer [illegible] type of solution as a function of transfer [illegible]. However, [illegible] measures of the learning process suggested some differences [illegible] II [illegible] [illegible] effects, [illegible] for him [illegible] measures of the learning process seem [illegible] difficulty-level [illegible] effects. The [illegible]. These differences detected by measures of the learning process in [illegible] way detract from the powerful effect of the cognitive transfer processes.

The hopes of obtaining specific data both on the exact time of transfer and the particular element or elements transferred did not materialize. Some Ss did not spontaneously verbalize such information, and among those who did, some admitted that they had been consciously transferring before verbalizing the fact. It is difficult to get an accurate time measure for Ss who do not spontaneously verbalize such information. After the experiment, some could only offer such vague information as "I was trying to do it before the first task-hint." And reports about the elements transferred were difficult to categorize because of their complexity. Though some Ss transferred only one element or combination of elements, others switched many times from one element or combination to another.

Nothing has been said about the unspecified associational mechanisms which cause uninstructed transfer hypotheses to occur. Though this question remains unanswered after this study the verbalizations of Ss suggest various possibilities. Many Ss without a transfer hint began to relate the tasks on the very first trial. Their transfer hypotheses might be related to a "suspicion" about the experiment or to something specific in the instructions for the transfer task. The second appearance of a circle and square or the second use of k and the same numbers for the k-scale might have been cues. If this is the case, calling the position values of figures in the second task "columns" should increase the amount of conscious transferring. Not using a circle and square and using a different response scale should decrease the amount of conscious transferring. Several Ss, including some given the transfer hint, abruptly began to transfer at some time during the transfer task. Those given the transfer hint apparently had forgotten it and then suddenly remembered it. Those not given the transfer hint frequently said that they had begun to wonder why they had

been given the first task. One S, while studying a card after feedback, mentioned several hypotheses, one of which was 2CS. Having said it aloud, she said "That's just like the first one," and she continued to use it. However, she was the exception, for other Ss actually solved with 2CS and failed to realize that is was identical to Rule 1 in the training task. One S, after 14 trials, said "Well, I guess I'll have to try the old formula." When questioned about this, he said he had thought from the beginning that the two tasks might be related, but preferred to solve the second task "on his own." A further study should be done to pinpoint these cues for transfer.

Getting students to use principles or knowledge in new situations seems to be an accepted educational goal. A transfer hint is clearly useful for this purpose given a goal of efficient learning in an immediate task. Whether giving transfer hints is good training for teaching students to transfer on their own is a different question. Perhaps in the long run the self-discovery method is better training for a habit of transferring. Over and above the fact of transferring or relating, Ss seemed to have more or less skill in finding the relationship when they set out to do so. Some Ss were unsuccessful even though they tried to find it. There seems to be a skill in transferring that needs to be developed over and above the habit of attempting to do so.

The high percentage of agreement between verbalized formulas and numerical answers was used as justification for analyzing the former as behavior which actually reflects cognitive processes. The 99.2% agreement was also supportive evidence for the theory of verbal control of overt performance presented by Dulany & O'Connell (1963). The fact that Ss occasionally used formulas which they thought had worked on the previous card been given the first task. One [illegible] trying a case after [illegible], mentioned [illegible] hypotheses, one of [illegible] 2CS. Having said it aloud, she said "this's just like the first one [illegible]" he continued to use it. However, she was the exception, [illegible] solved with 2CS and failed to realize that it was identical [illegible] the training task. One S, after 14 trials, said "[illegible], I [illegible] have to try the old formula." When questioned about this, he [illegible] the beginning that the two tasks might be related, [illegible] the second task "on his own." A further study should [illegible] these cues for transfer.

Getting students to use [illegible] in new [illegible] has been [illegible] accepted educational [illegible] for [illegible] purpose given [illegible] in [illegible] whether giving [illegible] for teaching transfer [illegible] the self-discovery method [illegible] transforming. Over and above the fact of [illegible] seems to have more or less skill in finding the relation [illegible] to do so. Some Ss were unsuccessful even though [illegible]. There seems to be a skill in transforming that [illegible] the habit of attempting to do so.

The high percentage of agreement between verbalized formulas and [illegible] was [illegible] for analyzing the former as behavior which actually reflects [illegible] processes. The [illegible] agreement was also supportive evidence for [illegible] verbal control of overt performance presented by [illegible] (19 [illegible]). The fact that Ss occasionally used formulas which [illegible] and worked on the previous [illegible].

but actually had not is data unrelated to this theory of verbal control. This theory states that simultaneous conscious processes control overt performance. It does not state that these processes must be accurate representations of past events. Memory, as everyone agrees, can be erroneous.

Given a complex task such as the transfer task, analysis of it in terms of a task model would seem to be a precarious venture. All analyses of process measures related to the task model militate against the strong assumption that _S_'s conception of the task is identical to _E_'s conception of it. Perhaps this gross discrepancy would not be true of a simpler task. but the assumption of identity of task models for _S_ and _E_ is very strong. The data regarding use of the known response scale and the use of prior information can be added to the evidence against strong assumptions. Subjects did not use the information that _k_ could only be whole numbers between 3 and 12 to any great extent. Only 29% of the verbalized formulas would have generated all and only the whole numbers on the _k_-scale. In fact, on 49 trials numerical answers that were off the _k_-scale were given. The _Ss_' use of prior feedback was also severely restricted. Of the verbalized formulas, 79% had been implicitly disconfirmed on the average of less than two trials (1.84 trials) prior. An additional 5% of the verbalized formulas had been explicitly disconfirmed by prior feedback. This lack of use of prior feedback suggest a severe memory problem and possibly the fact that _Ss_ cannot encompass more than the information of a very few trials into their hypothesis-testing. Furthermore, it is possible and likely that some _Ss_ do not even attempt to remember the _whole_ preceding card or cards. The amount of selective memory of this type that occurred cannot be assessed from the present data.

Feedback seemed to have three roles in this particular task. It was used to eliminate false hypotheses. This function was suggested by Harlow (1959). However, this function must be qualified by the fact that it functions over a short sequence of trials since Ss do use formulas which have been either explicitly or implicitly disconfirmed. Secondly, it was used as a basis for the generation of new hypotheses. Thirdly, it has a relationship to S's certainty. Though no measure of this function was obtained in this study, many Ss recognized that a formula could be fortuitously correct for a short sequence of trials. Each correct answer increased their certainty of having the correct formula. Though this was not necessarily true of all Ss, many were not sure that they had actually solved the transfer task after the four criterion trials. Additional criterion trials would have been needed to make them absolutely certain.

If these two tasks are viewed as the first two problems in developing a learning set, some suggestions about learning sets are possible. Ordinarily learning sets are viewed as some unspecified process of "learning how to learn." However, they could just as well be viewed as learning some set of specific things about a family of problems. One of these things is probably learning to relate all the problems since they are in fact related. Though this study gives only some evidence on this point, it would be worth pursuing.

Feedback seemed to have [illegible]

used to eliminate [illegible]

(1959). However, [illegible]

tion over a short sequence of trials [illegible]

either [illegible]

basis for [illegible]

for his certainty [illegible]

study, [illegible]

a short [illegible]

of having [illegible]

all Ss [illegible]

after the [illegible]

needed to [illegible]

If these [illegible]

a learning set, [illegible]

learning sets are viewed [illegible]

learn." However, [illegible]

specific things about [illegible]

learning to relate [illegible]

this study gives [illegible]

Bibliography

Adams, J. Laboratory studies of behavior without awareness. Psychological Bulletin, 1957, 54, 383-405.

Ausubel, D. P., Schpoont, S. H., and Cukier, Lillian. The influence of intention on the retention of school materials. Journal of Educational Psychology, 1957, 48, 87-92.

Azuma, H. Comparison of a correlational with a probabilistic approach to concept learning. Unpublished doctoral dissertation, University of Illinois, 1960.

Barker, R. G. Factors influencing transfer between finger mazes. Journal of General Psychology, 1932, 6, 115-132.

Cronbach, L. J. Educational psychology. New York: Harcourt, 1963.

Cronbach, L. J. & Azuma, H. Can we tell what the learner is thinking from his behavior? Paper read at Conference on Research Methodology in Training, Washington University, St. Louis, Missouri, 1961 (a).

Cronbach, L. J. & Azuma, H. Performance on a concept-attainment task with scaled cues. Unpublished manuscript, University of Illinois, 1961 (b).

Dulany, D. E., Jr. Hypotheses and habits in verbal "operant conditioning." Journal of Abnormal and Social Psychology, 1961, 63, 251-263.

Dulany, D. E., Jr. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In C. W. Eriksen (Ed.), Behavior and awareness: A symposium of research and interpretation. Durham, North Carolina: Duke University Press, 1962.

Dulany, D. E., Jr., & O'Connell, D. C. Does partial reinforcement dissociate verbal rules and the behavior they might be presumed to control? Journal of Verbal Learning and Verbal Behavior, 1963, 2, 361-372.

Deese, J. The psychology of learning. New York: McGraw-Hill, 1958.

Eriksen, C. W. Discrimination and learning without awareness: A methodological survey and evaluation. Psychological Review, 1960, 67, 279-300.

Gagne, R. M., & Smith, E. C., Jr. A study of the effects of verbalization on problem solving. Journal of Experimental Psychology, 1962, 63, 12-18.

Harlow, H. F. Learning set and error factor theory. In Koch, S. Psychology, a Study of a Science. Vol. 2, New York: McGraw-Hill, 1959, 492-537.

Bibl[illegible]

Adams, J. Laboratory studies of be[illegible] awareness. Psychological Bulletin, 1957, 54, 383-405.

Ausubel, D. P., Schpoont, S. H., & Cukier, [illegible]. The influence of intention on the retention of school materials. Journal of Educational Psychology, 1957, 48, 87-92.

Azuma, H. [illegible] with [illegible] approach to concept learning. [illegible] doctoral dissertation, University of Illinois, 1960.

Barker, A. G. [illegible] transfer between [illegible]. Journal of General Psychology, 1932, 6, [illegible].

Cronbach, L. J. Educational psychology. [illegible]: Harcourt, 1963.

Cronbach, L. J. & Azuma, H. [illegible] the learner is thinking [illegible] his behavior? [illegible] Research Methodology in Training, Washington University, St. Louis, Missouri, 1960 (a).

Cronbach, L. J. & Azuma, H. [illegible] concept-attainment task with scaled cues. Unpublished [illegible], University of Illinois, 1961 (b).

Dulany, D. E., Jr. Hypotheses [illegible] "operant conditioning." Journal of Abnormal and Social Psychology, 1961, 63, 251-263.

Dulany, D. E., Jr. The place of hyp[illegible] intention: An analysis of verbal control in verbal conditioning. In C. W. Eriksen (Ed.), Behavior and awareness: A symposium of research and interpretation. Durham, North Carolina: Duke [illegible], 1962.

Dulany, D. E., Jr., & O'Connell, [illegible] reinforcement [illegible] ciate verbal rules and the behavior [illegible] presumed to control? Journal of Verbal Learning and Verbal Behavior, 1963, 2, 361-372.

Deese, J. The psychology of learning. New York: McGraw-Hill, 1958.

Eriksen, C. W. [illegible] awareness: A methodological survey and evaluation. Psychological Review, 1960, 67, 279-300.

Gagne, R. M., & Smith, E. C., Jr. A study of the effects of verbalization on problem solving. Journal of Experimental Psychology, 1962, 63, 12-18.

Harlow, H. F. Learning [illegible] error factor theory. In Koch, S. Psychology: a Study of a Science. Vol. 2. New York: McGraw-Hill, 1959, [illegible].

Judd, C. H. The relation of special training and general intelligence. Educational Review, 1908, 36, 28-42.

Krasner, L. Studies of the conditioning of verbal behavior. Psychological Bulletin, 1958, 55, 148-171.

Mattson, D. E. Learning how to learn under several cue conditions. Unpublished doctoral dissertation, University of Illinois, 1963.

Osgood, C. E. Method and theory in experimental psychology. New York: Oxford University Press, 1953.

Paul, G., Eriksen, C. W. & Humphreys, L. G. Use of temperature stress with cool air reinforcement for human operant conditioning. Journal of Experimental Psychology, 1962, 64, 329-335.

Raven, J. C. & Lewis, H. K. Progressive Matrices (a non-language test of intelligence). New York: The Psycholgical Corporation, 1938, 1947, 1951.

Ruger, H. A. The psychology of efficiency. Archives of Psychology, 1910, 15, 88pp.

Salzinger, Kurt. Experimental manipulation of verbal behavior: A review. Journal of General Psychology, 1959, 61, 65-94.

Spielberger, C. E., Levin, S. M., & Shepard, Mary C. The effects of awareness and attitude toward the reinforcement on the operant conditioning of verbal behavior. Journal of Personality, 1962, 30, 106-121.

Stevenson, H. W., & Weir, M. W. The role of age and verbalization in probability-learning. American Journal of Psychology, 1963, 76, 299-305.

Stolurow, L. M., & McHale, T. J. Amount of information conveyed by a knowledge of either the principle or cues given in either algebraic or geometric form. (Data in process), 1965 (a).

Stolurow, L. M., & McHale, T. J. Some perceptual and verbal factors in the transfer from one task to another generated from the same model. (Data in process), 1965 (b).

Toda, M. Micro-structures of guess processes. Unpublished manuscript, Center for Cognitive Studies, Harvard University, 1962.